# Breast Density Estimation and Micro-Calcification Classification

**Minu George**

Department of Computer Science
Aberystwyth University

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

July 2021

I would like to dedicate this thesis to my loving parents and siblings who taught me to be trustworthy, honest and supportive, to my husband for being a guardian during my studies and to my little angel, Mannya for being my strength and inspiration to move forward. . . .

# Declaration and Statement

DECLARATION

This thesis has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed Minu George (candidate)

Date 09/07/2021

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Where **correction services**[1] have been used, the extent and nature of the correction is clearly marked in a footnote(s).
Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed Minu George (candidate)

Date 09/07/2021

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed Minu George (candidate)

Date 09/07/2021

[1] This refers to the extent to which the text has been corrected by others.

# Acknowledgements

I am deeply thankful to Lord God for his amazing grace, which has kept me from the beginning of my life until now. I would like to express my gratitude to my supervisor Prof. Reyer Zwiggelaar for the guidance and support throughout my PhD studies. Especially for his moral support, advice for managing a balance between work and family during stressful times. I feel lucky to be his student as he made me confident, motivated and comfortable in work during my bad times to complete the studies in time. He was a great role model as a teacher and supervisor from whom I learned a lot of good characters apart from research studies. So I thank him for all his time, support and encouragement during the entire PhD study.

I want to thank Dr. Bernie Tiddeman for his support, assistance, effort and time in finalizing the thesis draft. My acknowledgement also goes to Dr. Zhili Chen for her support, time and suggestions in improving the thesis work.

I want to thank all the people from my research lab, Vision, Graphics and Visualisation Group for their helpful feedback, questions, and suggestions on my research during meetings. I want to thank my colleagues, especially to Andrik, Liping, Zobia, Azam, Nashid and Ahmad for the supportive environment in my hard times.

I acknowledge all the staff members of the Computer Science department, especially from the Computer Vision group for supportive suggestions and comments during the presentation and other technical supports when needed.

This acknowledgement would not be complete without thanking my family, especially my husband Tibin, for his understanding, encouragement, support during hard times. The acknowledgement will be meaningless without thanking my little angel, Mannya Isa (Vavachi) for all her supports and smiles. You were the inspiration, and you made me stronger to face all hurdles and complete work in time. I love you to the moon and back. I want to thank my parents for all their love, blessings, prayers and for taking care of my baby to complete my studies.

# Abstract

Breast cancer is the most common cancer among women worldwide. Mammography is considered as the golden standard and effective screening tool in detecting breast cancer at its early stage. The extensive screening programmes and health care standards lead to the necessity of Computer-Aided Diagnosis tools to support radiologists. In this thesis, we investigate the potential of computer vision and image processing algorithms in developing CAD algorithms for breast density estimation and micro-calcification classification. Many studies have shown that breast density and parenchymal patterns are reliable indicators for breast cancer risk. So, for risk assessment, the thesis proposed a multi-scale elliptical blob modelling for parenchymal pattern representation based on a statistical analysis of breast tissue. To describe the breast density type and risk, we analysed the breast for blob-like dense tissue. The study showed a relation with the BI-RADS density classification based on fatty and dense tissue blob structures. Subsequently, the thesis proposed a new variant to LBP called Mean-Elliptical Local Binary Pattern (M-ELBP), where the intensity and texture features were combined in an elliptical topological structure. The method is more robust to noise as it does not perform a direct comparison between pixels as in traditional LBP and has the benefit of extracting features from multiple orientations. The studies based on density estimation investigated the potential of ROI size, descriptor size, and classifier effect on density classification. The validity of the proposed methods is evaluated using the MIAS and DDSM databases. The obtained classification accuracy was up to 77.4% for MIAS tissue-based class. BI-RADS based classification attained up to 75.4% and 47.65% for MIAS and DDSM. For microcalcification classification, a novel method called connected-chain is developed based on analysis of the topological/distributional pattern of micro-calcifications in MCCs. The connectivity/closeness between individual micro-calcification was analysed to discriminate between benign and malignant case. The performance of the method was tested using three datasets: MIAS, DDSM, OPTIMAM obtaining classification accuracies of 82.50%, 86.47% and 76.75%, respectively.

# Table of contents

# List of figures

# List of tables

# Chapter 1

# Introduction

The chapter provides the context for the thesis by delivering an insight into the breast cancer status in Section 1.1 followed by different breast imaging modalities and breast anatomy in Section 1.2. Risk assessment is an essential aspect of controlling the breast cancer mortality rate. To address this, Section 1.3 presents different breast cancer risk factors and risk assessment methods. From existing research studies, breast density is identified as an essential risk factor in developing breast cancer, to take this into account, Section 1.4 discusses different breast density classification schemes. The detection and classification of mammographic abnormalities are an essential stage in Computer-Aided Diagnosis (CAD) systems; therefore, mammographic micro-calcifications are explained in Section 1.5 followed by research aim and contributions in Section 1.6 and finally thesis outline in Section 1.7.

## 1.1 Breast Cancer Status

According to the World Health Organization (WHO) survey, breast cancer is the most common and frequent cancer among women accounting for 11.6% of new cases world-wide (2.09 million cases in 2018). Almost 2.1 million women are affected by breast cancer each year, and it accounts for the most significant number of cancer-related deaths among women. It is estimated that in 2018, almost 627,000 women died of breast cancer which is nearly 15% of all cancer deaths among women. Approximately 70% of deaths due to disease occur in low- and middle- income countries due to the lack of screening, failure in early detection, and absence of awareness programmes. As from the Globocan 2018 project by WHO, incidence rates are high in developed countries compared to Asia and Africa as shown in Figure 1.1, however, mortality rates were comparatively low compared to developing countries. In the US, as per

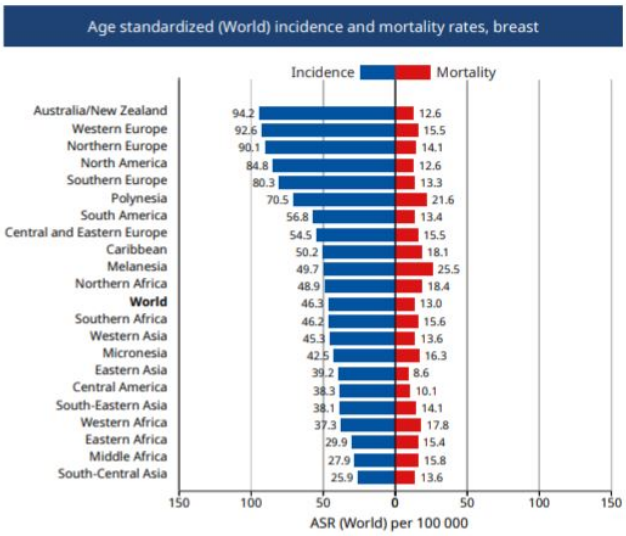



Fig. 1.1 Breast cancer worldwide incidence - mortality rate by Globocan 2018 report

studies, 1 in 8 women tend to develop breast cancer in their lifespan (i.e. about 12.4%). An estimate of 266,120 new cases of invasive breast cancer is expected to be diagnosed this year. However, the death rates have been decreasing since 1989, especially for women above 50 years due to advanced treatments, screening and increased awareness like reducing the use of hormone replacement therapy (HRT) and changes in lifestyle (US breast cancer statistics).

Similarly, in the UK, the NHS Breast Screening Programme has invited all women aged between 50 and 70 for screening every three years. According to studies, breast screening has reduced the number of deaths from breast cancer by about 1,300 a year through early detection and thus improved the survival rate by almost five years. Survival of a breast cancer patient depends on many factors like the stage of cancer detection, age, size of the tumour, and treatment facilities (da Costa Vieira et al., 2017). The incidence of breast cancer has been lower in developing countries when compared to developed countries, but they tend to have a higher mortality ratio (Parkin et al., 2005). Since healthcare resources like screening programmes have been less developed in developing countries, a direct comparison of the incidence rate is less meaningful.

## 1.2 Breast Imaging Modalities

Breast cancer detected at an early stage has been easier to treat and has shown a higher chance of survival rate. According to a cancer research UK survey, more than 90% of women who were diagnosed at an early stage survive for at least five years than

compared to 15% women who got diagnosed at an advanced stage of the disease. Similarly, studies from the US show that breast cancer mortality has decreased to 34% from 1975 to 2010 due to early detection of cancer through screening and advanced treatments (Narod et al., 2015). Moreover, recent studies in Ireland and Switzerland have shown that with testing and advanced treatments, breast cancer mortality rates are decreasing compared to previous years (Hanley et al., 2017; Herrmann et al., 2018; Mayor, 2017). Similarly, early detection of breast cancer in young women could increase the survival rate and reduce the chance of more invasive cancer (Elmore et al., 2005; Kadaoui et al., 2012).

Given such circumstances, the early diagnosis of breast cancer is vital and crucial in reducing the mortality rate. As imaging modality standards improve, the purpose of each modality will improve continuously reaching the goal of reducing breast mortality rate through early detection and treatment of invasive cancer (Autier and Boniol, 2018; Prasad et al., 2007).

### 1.2.1 Breast Anatomy

For studying the causes of breast cancer, knowledge of breast anatomy is essential. Breast tissue is overlying the chest (pectoral) muscle. Various tissue structures determine the size, shape, and texture of the breast, causing variations according to age and hormonal changes. The breast tissue structure consists of glandular tissue, fibrous tissue connecting lobes, and fatty tissue in intervals between the lobes. The breast contains approximately 15 to 20 lobes of glandular tissue constituting parenchyma of the mammary glands. They help in providing characteristic shape to the breast due to the arrangement of fatty tissue between lobes, and histological studies have shown that they are composed of lobules connected by areolar tissue, ducts and blood vessels (Tobon and Salazar, 1975; Zucca-Matthes et al., 2016).

The adipose tissue of the breast lies typically between lobes rather than within lobules (Zucca-Matthes et al., 2016). The breast also contains blood vessels, lymph vessels, and lymph nodes. The lobules produce milk and are generally called glandular tissue. The breast appears to be dense if the amount of fibrous or glandular tissue is considerably more than the fatty portion. Breast density decreases with age. An overall structure of breast tissue is shown in Figure 1.2. Parenchymal patterns in breast imaging are essential as it affects mammographic screening sensitivity and explains the density of the breast. A detailed parenchymal structure is shown in Figure 1.3.

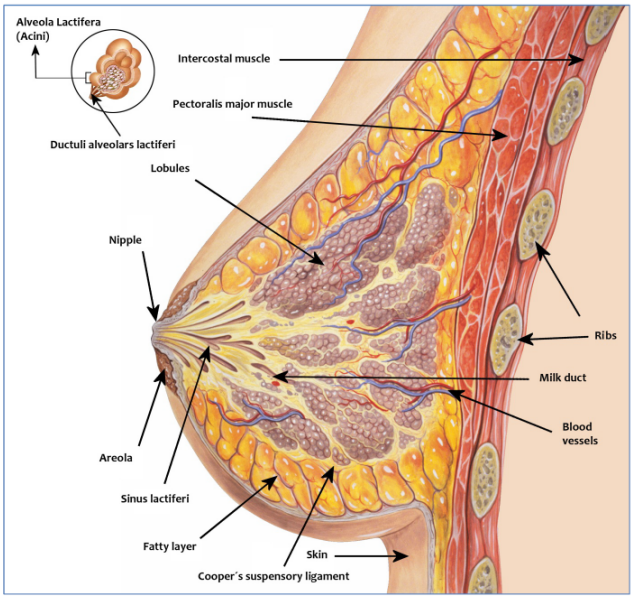



Fig. 1.2 Breast anatomy showing the various tissue and internal structure (the image taken from https://studylib.net/doc/8232709/anatomy-of-the-breast).

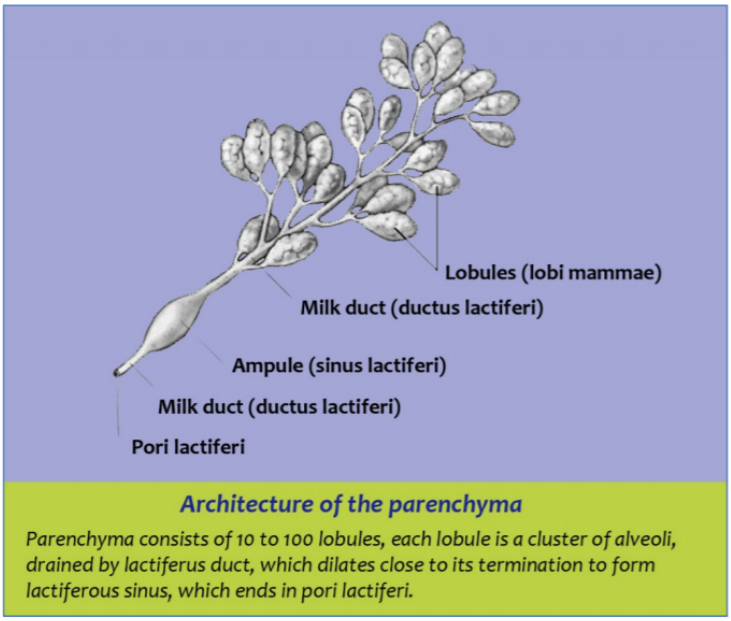



Fig. 1.3 Detailed representation of a parenchymal pattern (image taken from https://studylib.net/doc/8232709/anatomy-of-the-breast).

### 1.2.2 Mammography

From an imaging perspective to diagnose breast cancer, mammography has been considered as the "gold standard", despite debates on the effectiveness of mammographic screening programmes in checking breast cancer mortality ratio (Prasad et al., 2007). Mammography has been recognised as a primary imaging modality and most effective tool for early-stage breast cancer diagnosis (Feig et al., 1977; Greif, 2010; Kolb et al., 2002; Tabar et al., 2003) as it could detect lesions even before they become palpable and physically examined (Alanís-Reyes et al., 2012). Nevertheless, debates are still going on with the risk of mammographic screening process on topics like effectiveness on controlling mortality rate (Gøtzsche et al., 2009), overdiagnosis (De Gelder et al., 2011), and false positive (Brewer et al., 2007). Besides, other risks like radiation through continuous screening (Feig and Hendrick, 1997) also exists, however, mammography is believed to be able to diagnose breast cancer at an early stage and thus saving lives by preventing advanced cancer stages (Welch and Frankel, 2011). To this end, the mammographic screening process has shown to diagnose early-stage breast cancer (Lee et al., 2010; Qaseem et al., 2007).

Mammography is a breast imaging process where low-dose x-rays are used to detect cancer. The x-ray/radiograph is a noninvasive medical test that helps radiologists to examine the internal structure of breast for any abnormalities. The breast region is exposed to low dose ionising radiation (30 kVp) to generate images of the internal breast tissue structure (the images are known as maommograms). X-rays are considered the most frequently used and oldest form of medical imaging. The procedure is similar to other x-rays but with a low dose, presenting a high-quality image with good contrast, resolution, and low noise (Sivaramakrishna and Gordon, 1997).

A mammography unit produces x-rays and has a specialised frame that allows only exposure of the breast region to x-rays. These units help in holding, compressing, and positioning the breast so that images can be obtained at different angles for a more detailed examination. A model of a mammography unit is shown in Figure 1.4.

An x-ray machine produces a burst of x-rays to pass through the breast region to a detector placed on the opposite side. The sensor can be a photographic film plate (which captures the x-ray image on a film) or a solid-state detector (transmits the electronic signals to a computer in the form of digital data). The images produced are called mammograms representing the internal patterns of the breast. For example, low-density tissues like fat appear translucent (darker shades of grey) and areas of dense tissue such as connective tissue, glandular tissue, tumours, masses seem to have

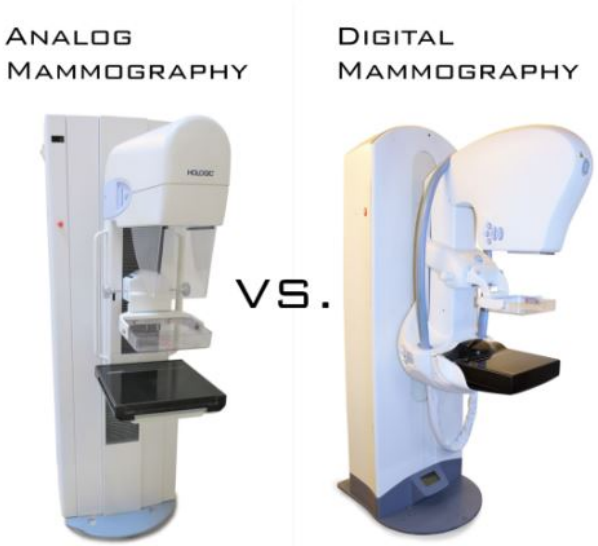



Fig. 1.4 Analog vs Digital Mammography Machines (the image was taken from https://lbnmedical.com/wp-content/uploads/2017/12/anvsdig-3.png).

higher intensity in mammograms. Breast compression is essential for getting useful mammograms (Kopans 3rd, 2007) as it prevents movement of the patient, avoids superimposition of tissues to an extent providing a better lesion view, less radiation is passed due to reduced thickness. Besides, it helps in spreading out breast tissue uniformly giving a better understanding of tumorous tissue.

### Abnormality reading from Mammogram

The areas which do not appear to be normal tissue are a concern for a radiologist. Radiologists look for higher intensity areas, highly dense tissue, note the size, shape, pattern, and edge of each suspected abnormality. Possible abnormalities include:

- **Cysts:** Which are small-sized fluid-filled sacs with a thin wall and may not be cancerous.
- **Calcifications:** Which are small deposits of calcium in breast tissue that appear as white dots in a mammogram. Depending on the size of the deposit, it is called macrocalcification (could be due to ageing and are generally benign) or microcalcifications (could be cancerous). Microcalcifications requires further investigation by a radiologist depending on their pattern. Figure 1.5 shows an example of microcalcifications.

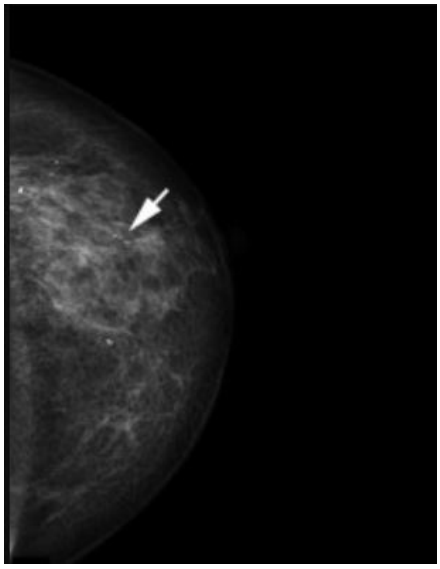
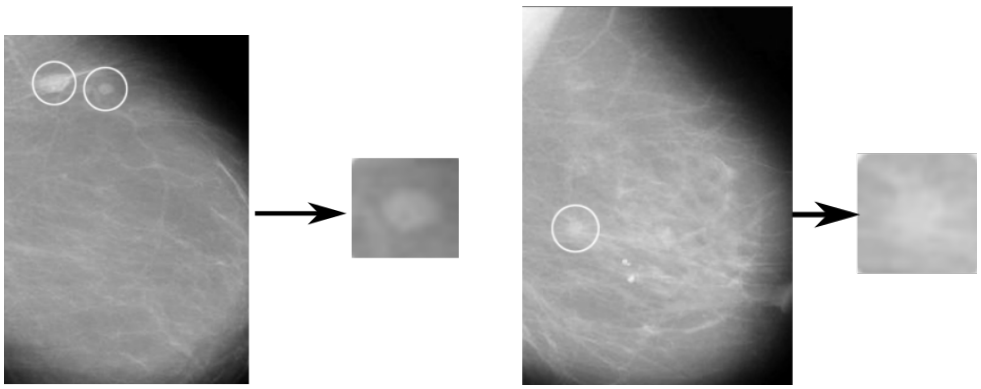


Fig. 1.5 Coarse heterogeneous microcalcifications in cranio-caudal (CC) view (the image was taken from Henrot et al. (2014)).

- **Fibroadenomas:** Which are usually benign tumors which feel like marble and common at any age.
- **Scar:** Scar tissue is caused by scarring on breast tissue and appears as high intensity on a mammogram.
- **Masses:** A mass is defined as a space-occupying lesion, visible in two different projections, characteristic by its shape and contour in a mammogram seen as a bright dense region as shown in Figure 1.6 (Berment et al., 2014).
- **Architectural distortion:** The Breast Imaging Reporting and Data System (BI-RADS) defines an architectural distortion as an appearance in which "the normal architecture of the breast is distorted with no definite mass visible, including spiculations radiating from a point and focal retraction or distortion at the edge of the parenchyma. Architectural distortion can also be an associated finding" (Gaur et al., 2013). Examples are shown in Figure 1.7.

Mammograms also provide the information about the density of breast tissue; women with more dense tissue have a higher risk of developing cancer (Zdon et al., 2019) and also it is more difficult for radiologists to localize abnormalities as density can hide irregularities (Lattanzio and Guerrieri, 2018).



(a) (b)

Fig. 1.6 Examples of mammographic masses (a) Mammogram with a circumscribed mass (b). Mammogram with a spiculated mass (figure was taken from Özekes et al. (2005)).

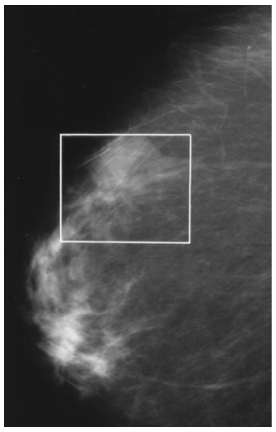


Fig. 1.7 Architectural distortion (image was taken from (Baker et al., 2003)).

### Mammographic Views

In mammographic screening, each breast is imaged separately at two to four angles: (a) Medio-Lateral Oblique (MLO) view (oblique or angled view), (b) Cranial Caudal (CC) view (view from above), (c) mediolateral (ML) view (view from the centre of the chest outward) and (d) latero-medial (LM) view (view from the outer side of the breast, towards the middle of the chest). Radiologists prefer MLO and CC view with proper positioning for screening both breasts getting a total of 4 images per patient (Bassett et al., 1993; Eklund, 2000; Popli et al., 2014; Sickles, 1988). The CC view can demonstrate maximum tissue on both medial and lateral aspects of the breast with retro mammary space and some pectoral muscle (Popli et al., 2014) as represented in Figure 1.8. MLO view demonstrates axilla, axillary tail, and infra-mammary fold with all the breast tissue as in Figure 1.9.

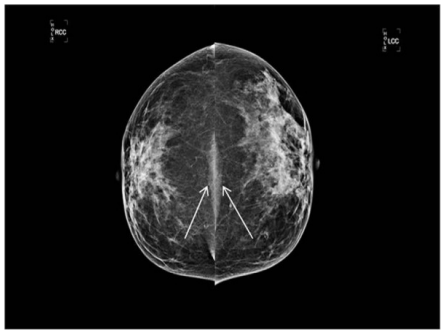


Fig. 1.8 A CC view with retromammary space and pectoralis muscle (arrows). Visualization of pectoralis muscle on CC view implies that no breast tissue along chest wall has been excluded (the image was taken from Popli et al. (2014)).

Fig. 1.9 Bilateral MLO: pectoralis muscle forming "V, " when viewed as mirror images (the image was taken from Popli et al. (2014)).

## 1.3 Breast Cancer Risk Factors and Risk Assessment

Irrespective of many studies on breast cancer risk factors, the real causes for developing breast cancer are unknown (Eccles et al., 2013; Kamangar et al., 2006; Torre et al., 2015). Risk factors causing breast cancer can be of two types: (a) inherent factors such as age, sex, race, genetic makeup and other changes to patterns in breast tissue during the lifetime, (b) extrinsic type like lifestyle, diet or long-term medical intervention, hormonal treatments (Kamińska et al., 2015). Often age is a significant risk factor in developing breast cancer, especially for women above 45-50 years (Ataollahi et al., 2015; Jemal et al., 2009; Steiner et al., 2008). Similarly, inherited genes like BRCA1 (BReast CAncer gene one) and BRCA2 (BReast CAncer gene two) are the genes linked with breast cancer and ovarian cancer if they function abnormally (Brody and Biesecker, 1998). There are different models for assessing breast cancer risk, which

include multiple factors (Singletary, 2003). The most widely implemented non-image based risk models are explained below:

**Gail Model:** Widely used model for general risk factors. Gail et al. (1989) assessed a variety of potential risk factors for breast cancer using unstratified logistic regression analysis. They analyzed data collected from 2,852 white women with breast cancer and 3,146 white women controls who underwent breast screening every year. They found that certain expected factors like cigarette smoking, use of oral contraceptives, or alcohol consumption were not significant risk factors for breast cancer. Instead, significant determinants of risk in this population of women were: (1) family history in a first-degree relative, (2) late age at childbirth, (3) early menarche, and (4) multiple previous benign breast biopsies. The major drawback of this analysis was limited information regarding family history. Moreover, analysis ignored second-degree relatives, age at which cancer developed for relatives, and findings of ovarian cancer or lobular carcinoma in situ.

**Claus Model:** To address some deficits of the Gail model, the Claus model is used. It analyzed data from 4,730 histologically confirmed breast cancer patients aged 20 to 54 with and 4,688 controls matched for age (by 5-year age category) and geographic location (Claus et al., 1994). The primary risk assessment was based on age, number, and type of relatives who were affected. But this model has the disadvantage that it is appropriate only for particularly high-risk patients who have at least one female relative who is diagnosed with breast cancer.

Other models like **BRCAPRO** considered the probability of family history of first- and second- degree relatives related to a mutation in a BRCA gene (Berry et al., 1997) like whether they have been diagnosed by unilateral, bilateral or ovarian cancer in their lifespan. Similarly, the **Bodian** model calculates the risk of invasive or in-situ breast cancer based on Lobular carcinoma in situ (LCIS) and age at detection (Bodian et al., 1996).

In recent years, there has been growing interest in segmenting, classifying the breast density and parenchymal patterns as risk factors forming the primary focus of research studies. Various studies based on investigating the structure of parenchymal patterns and breast density indicated a high risk for breast cancer in highly dense breasts (Lattanzio and Guerrieri, 2018; Wolfe, 1976a; Zdon et al., 2019). Incorporating image-based risk factors with existing non-image based models could give a better understanding of the risk factors for controlling breast cancer.

### 1.3.1 Mammographic Risk Assessment

Mammographic risk assessment provides an indication of a woman's risk in developing breast cancer based on imaging modalities like mammograms assisting early detection and treatment. Mammographic parenchymal patterns and density are both reliable predictive image-based indicators for breast cancer risk analysis based on mammographic features. Breast density is the measure of radiodense fibroglandular tissue in the breast (McCormack and dos Santos Silva, 2006). Extensive breast density and parenchymal patterns can affect mammographic sensitivity. Besides, it might increase the chance of missing abnormalities hidden in dense tissue obstructing cancer detection at an early stage (Carney et al., 2003; Kerlikowske et al., 1966; Mandelson et al., 2000; Rosenberg et al., 1998). In a mammographic image, malignant tissue has negligible physical density difference from fibrous breast tissue or dense tissue resulting in low contrast. So there are chances of missing carcinoma and microcalcification during diagnosing time due to similar contrast between tissue (Boyd et al., 1998; Brisson, 1991; De Stavola et al., 1990; He et al., 2015; Saftlas and Szklo, 1987; Titus-Ernstoff et al., 2006; Zdon et al., 2019).

The majority of studies investigating breast density assume that breast tissue is either adipose or glandular. The glandular tissue constitutes supportive tissue, connective tissue (stroma), ductal structures, and epithelium tissue. On the other hand, some studies focus on the complexities of the parenchymal patterns for breast density classification. The mammographic appearance of the breast varies among women and depends upon the breast tissue composition and X-ray attenuation properties of various tissue types, as mentioned in Section 1.2.2. Non-fatty tissue like connective and epithelium tissue is radiographically dense compared to fatty tissue. It would absorb more X-rays during the image acquisition process, appearing to be brighter in mammographic images (Boyd et al., 1998, 2007; Wolfe, 1976a). The different mammographic appearances generated by variations in tissue composition are referred to as mammographic parenchymal patterns. They represent morphology, composition, and spatial distribution of various breast tissue. From an image processing view, bright regions formed by non-fatty tissue are referred to as mammographic density, which in turn reflects the relative amount of dense tissue composition in the breast.

Section 1.4 describes classification schemes used to define mammographic density and how they have been employed in clinical studies. The main classification schemes used by radiologists for mammographic risk assessment are Wolfe, Boyd, BI-RADS, and Tabar. This classification is based on either parenchymal patterns or breast density depending on which plays a more significant role in assessing the development of breast

cancer risk. While Boyd et al. (1995) and BI-RADS (D'Orsi, 2013) classification schemes focus on mammographic density, Wolfe (1976a) and Tabár and Dean (1982) focus on complexity of parenchymal patterns for estimating risk of breast cancer though they all are mutually correlated in risk assessment.

## 1.4 Breast Density Classification Categories

Breast density classification schemes could be quantitative or qualitative depending on the studies. Quantitative studies like Boyd's classification scheme helps in explaining the strong association between breast density and risk as they are less subjective (Boyd et al., 1982). However, a qualitative scheme helps in giving information on whether a particular woman with a high-risk parenchymal pattern structure would benefit from further follow-up on imaging modalities like ultrasound or tomosynthesis.

### 1.4.1 Wolfe Four Classification Scheme

Wolfe (1976a) considered parenchyma as composed of three building components: fat, connective and epithelial tissues, and "prominent ducts". Wolfe (1976a) considered connective and epithelial tissue together as 'mammary dysplasia'. Based on this, Wolfe categorizes mammograms into four parenchymal patterns as follows:

1. N1 (primarily fatty, lowest risk): mammogram is composed of fat (N = normal) and no amounts of dysplasia or a few fibrous tissue strands.

2. P1: this pattern includes fat as well as linear densities (enlarged ducts) occupying no more than 25% of the breast.

3. P2: linear densities (from enlarged ducts) occupying more than 25% of the breast. They are prominently in the upper outer quadrant but may be distributed throughout the breast (P = prominent ducts)

4. DY: dense, radiopaque breast (Dy=dysplasia); these patterns are again subdivided into low-risk (N1 and P1) and high-risk (P2 and DY) patterns (Wolfe, 1976b).

A fifth category was later added by Wolfe to these four:

Qdy (quasi-dysplasia): this group consists of young women whose dense breasts have a somewhat spongy texture due to fatty infiltration.

### 1.4.2 Boyd's Six Class Categories

Boyd et al. (1995) and colleagues performed a quantitative study (in 1980) as they felt Wolfe's classification had the limitation of a subjective scheme of examination. They provided a more explicit categorization of dysplasia than defining DY as severe. So, based on the proportion of the breast occupied by mammographically dense tissue (estimate the proportion of fibroglandular tissue), a scheme was developed called Boyd Six Category Classification (SCC). Based on mammographic density percentage given by radiologists and divided into six categories of unequal intervals:-

1. SCC1: 0%,
2. SCC2: $> 0 - 10\%$,
3. SCC3: $> 10 - 25\%$,
4. SCC4: $> 25 - 50\%$,
5. SCC5: $> 50 - 75\%$,
6. SCC6: $> 75\%$.

### 1.4.3 Tabár Five Classification Patterns

Tabár and Dean (1982) classified mammographic images into five patterns based on parenchyma, rather than the proportion of mammographic density, and similar to Wolfe (1976a). They suggest that the model of breast tissue in mammographic images contain four mammographic building blocks: nodular densities (N), linear densities (L), homogeneous fibrous tissue (H), and radiolucent fat tissue (R). Mammograms were subdivided into five risk categories, based on the distribution of building blocks [N%, L%, H%, R%] (Tabár et al., 2004).

1. Pattern I: [25%, 15%, 35%, 25%], A balanced proportion of all components and is considered as lowest risk pattern of breast tissue with a slight predominance of fibrous tissue.
2. Pattern II: [2%, 14%, 2%, 82%], predominance of radiolucent fat tissue (fat breast) providing radiologist excellent environment in detecting abnormalities.
3. Pattern III: [similar to Pattern II], the predominance of fat tissue with retro areolar residual fibrous tissue.

4. Pattern IV: [49%, 19%, 15%, 17%], predominantly nodular densities and appears to be resistant to the process of involution, making the perception of pathological lesions difficult on mammograms.

5. Pattern V: [2%, 2 89%, 7%], predominantly fibrous tissue (dense breast) with the highest risk. The mammogram is dominated by a voluminous amount of homogeneous structure-less fibrous tissue. This will form a ground glass-like appearance with no perceptible features making it difficult for radiologists to reveal small pathological lesions and abnormalities.

### 1.4.4 BI-RADS Classification of Density

BI-RADS (Breast Imaging Reporting and Data System) is a quantitative method established by the American College of Radiology (ACR) (D'orsi et al., 1998). It was designed as a quality assurance tool for mammographic reporting. Moreover, it provides insight to physicians about decreased mammographic sensitivity in detecting cancer with increased breast density (Sickles, 2007). It is presently the most widely used mammographic breast density classification method in North America and Europe (Eberl et al., 2006; Orel et al., 1999; Taylor et al., 2011). Fig. 1.10 shows example mammograms for each class.

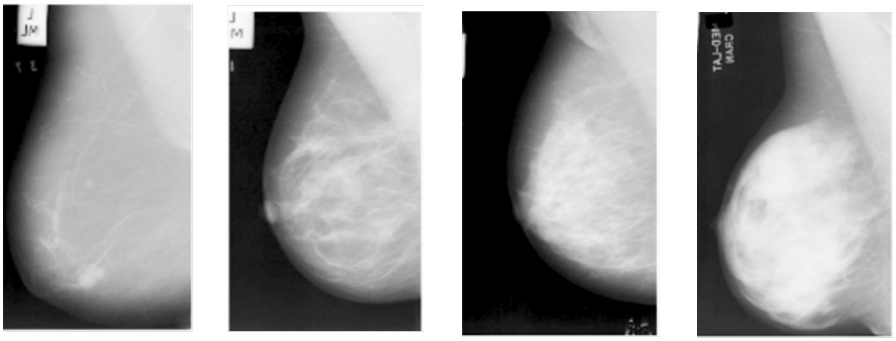



Fig. 1.10 Example mammograms showing various BI-RADS density classification based on BI-RADS $5^{th}$ edition: (a) BI-RADS 1 (entirely fatty breast tissue), (b) BI-RADS 2 (scattered fibro-glandular density), (c) BI-RADS 3 (heterogeneously dense breast obscuring small masses), and (d) BI-RADS 4 (extremely dense breast lowering the sensitivity of mammography).

1. Category A / (BI-RADS I): The breasts are almost entirely fatty.

2. Category B / (BI-RADS II): There are areas of scattered fibroglandular density.

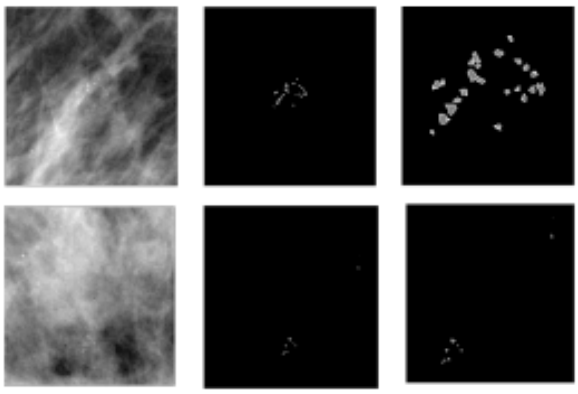


Fig. 1.11 Examples of ROI patches of malignant (top row) and benign (bottom row) microcalcification from mammograms, first column: Original patch; second column: annotated microcalcifications; third column: zoomed microcalcifications

3. Category C / (BI-RADS III): The breasts are heterogeneously dense, which may obscure small masses.

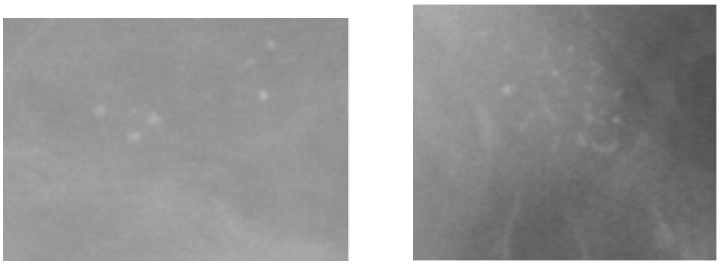
4. Category D / (BI-RADS IV): The breasts are extremely dense, which lowers the sensitivity of mammography.

Muhimmah et al. (2006) investigated the correlation between the four classification schemes (i.e. Wolfe, Boyd, BI-RADS, and Tabár) using the MIAS (Mammographic Image Analysis Society) database (Suckling et al., 1994) and found that there was a strong correlation between Wolfe, Boyd, and BI-RADS categories. Since it is considered as the standard evaluation method worldwide, we chose to use BI-RADS classification in our studies.

## 1.5 Microcalcifications in Mammograms

Breast microcalcifications are small spots of calcium deposit which are common among women and are mostly benign (Cheng et al., 2003). Microcalcifications appear as white specks on mammograms, as shown in Figure 1.11. However, the presence of fine, patterned granular microcalcifications clusters can be an indication of early breast carcinoma and of concern to radiologists leading to further histological examinations for clarifications.

One of the prominent features of microcalcifications is that they are found in a group or clusters. Microcalcification clusters (MC) are defined as a true cluster if there are more than 3 microcalcifications in a $1cm^2$ region of the mammogram (Hernández et al., 2016; Ma et al., 2010; Soltanian-Zadeh et al., 2004). The classification



(a) (b)

Fig. 1.12 ROI samples for benign (a) and malignant (b) micro-calcification clusters.

of microcalcification clusters as benign or malignant is more difficult compared to masses (Cheng et al., 2003), which leads to one of the aims of this thesis (see section 1.6). From various researches and clinical studies, it has been predicted that benign microcalcifications clusters have fewer microcalcifications than malignant clusters. Similarly, they appear as coarse, round or oval shapes with a uniform shape and size with high density, and normally homogeneous with scattered and diffuse patterns (Mousa et al., 2005). It is stated that malignant microcalcifications have more irregular and dissimilar appearance than benign cases with size and shape variance, forming microscopic and fine, linear branching patterns. Besides, they tend to have a more grouped pattern than the diffuse pattern of benign micro-calcifications (Feig et al., 1987; Sickles, 1986). Figure 1.12(a) shows an example of a benign and Figure 1.12(b) a malignant micro-calcifications cluster.

## 1.6 Research Aims and Objectives

As seen from breast cancer statistics, breast cancer is one of the significant cancers among women worldwide. Therefore, any research contribution to controlling the mortality rate due to breast cancer is vital.

According to Section 1.3, parenchymal breast patterns, and breast density are major risk factors in developing breast cancer. Therefore, one of the focus areas of the thesis is on modelling breast parenchymal patterns to understand the breast anatomical structures using geometrical structures from an image analysis perspective. In the current work, we attempt to model parenchymal patterns using elliptical blob-like structures for representing various mammographic patterns in various orientations. The study helps towards a better understanding of the complex nature of breast tissue patterns that can be used for risk assessment.

Considering breast density as one of the significant risk factors in developing breast cancer, we try to estimate breast density through mammographic tissue analysis using various texture analysis methods. To understand the importance of intensity of mammographic tissue in addition to texture in density estimation, we develop a new CAD algorithm called Mean-Elliptical Local Binary Pattern (M-ELBP) for density estimation. An attempt to investigate the effect of size and location of ROIs in density estimation is performed, in addition to considering the descriptor size and classifier size in density estimation.

Section 1.5 explains the role of abnormalities (micro-calcification) in breast cancer diagnosis. Therefore, another focus of our research is on developing a CAD algorithm for classification of segmented micro-calcification as benign or malignant. The study emphasizes the distributional pattern of MC in MCC. Based on the distributional arrangement, we propose a method called the connected-chain model for MC classification. Moreover, we investigated the applicability of the connected chain method for surrounding tissue texture analysis.

To this end, the main aim of the thesis is to analyze mammographic images to develop a framework for (a) breast cancer risk estimation, (b) to develop a model for breast parenchymal patterns as well as (c) develop a new method for the classification of microcalcification clusters using image processing and computer vision techniques.

In light of the above discussion, we subdivide the main goals into a set of objectives as mentioned below:

- To develop a method for breast density estimation. In addition, investigate the effect of ROI size and location, descriptor size, and classifier effect on breast density estimation along with a comparative study with related algorithms in breast density estimation.

- To develop a model for the representation of parenchymal patterns in mammographic images using elliptical blob-like structures. The bright elliptical blob-like tissue structures are represented as multiscale elliptical blobs at different orientations.

- Finally, micro-calcification distributional patterns based on distance were explored in developing a new algorithm called the connected-chain model for MC classification as benign or malignant. The proposed algorithm was compared and combined with texture features of surrounding tissue to estimate the potential of the method.

## 1.7 Thesis Outline

This section describes the organization and the flow of the thesis. As discussed, the thesis is focused on methods for mammographic breast density estimation, breast parenchymal modelling to estimate risk and chance of developing breast cancer in women from mammogram images. Besides, a method for classification of micro-calcifications is developed to distinguish benign and malignant micro-calcification abnormality.

Chapter 2 describes the need for CAD in medical image analysis for controlling cancer, followed by different concepts and methods used in the research for breast density estimation and risk scoring. The next section explains the different BI-RADS classification descriptors for micro-calcification detection, followed by MC abnormality classification. In addition, different texture analysis methods used in computer vision for image processing are explained to define the use of different texture analysis methods for feature extraction in breast density estimation and also in MC classification. The chapter gives a detailed study of different texture analysis techniques like statistical, structural, signal processing, and model-based methods. The next section provides a brief description of the uses of texture analysis like classification, segmentation, synthesis, and shape analysis. It explains a few texture feature descriptors used in our study for mammographic breast density estimation and micro-calcification classification. In the last section, the datasets used in the current study is presented.

Chapter 3 explains the different variants of Local Binary Patterns (LBP) and its applications. The proposed Mean Elliptical Local Binary Pattern (M-ELBP) variant for breast density estimation is introduced in the following section. The method is used to classify mammographic images into different BI-RADS classes. The last section describes the use of the proposed method for MIAS 3-class tissue classification along with a comparative study between different variants of LBP, the effect of choosing ROI size and location. Additionally, the effect of descriptor size and role of classifier in classification is studied. The study found that M-ELBP performs better than many variants and traditional LBP.

Chapter 4 introduces the proposed elliptical blob analysis model for breast density estimation and risk assessment. The chapter gives a brief description of different synthetic and simulated models used for breast parenchymal modelling and its applications in CADs. The chapter explains how mammographic structures are represented in scale-space and how they can be used for risk assessment. Moreover, the chapter introduces the elliptical blob modelling of mammographic patterns in scale-space at

different orientations. The investigation describes the potential of modelling breast tissue according to Tabar's model.

Chapter 5 introduces a new method for micro-calcification classification. The new approach, called connected chain modelling, is developed based on the distributional pattern of MCs in the micro-calcification cluster (MCC). Similarly, the chapter describes the multi-scale approach for connected chain modelling using morphology. Additionally, a comparative study based on texture features from the surrounding tissue of MCC and the connected-chain model is presented.

Chapter 6 gives a summary and conclusion of the thesis. Moreover, the novel contribution from the thesis and insight into possible future works are discussed. Lastly, the publications from the research are provided.

## 1.8 Summary

The chapter gave a vision of the status of breast cancer and why it is essential to develop more CAD algorithms to control breast cancer. The chapter explained breast anatomy, representation of breast tissue in mammography procedure, and possible abnormalities that could be detected from a mammographic image. Besides, the different risk assessment models and breast density categories for risk and density estimation were introduced. Importantly, the chapter provides an insight into the research aims and objectives of the thesis.

# Chapter 2

# Literature Review

## 2.1 Overview

In the previous chapter, we have discussed statistics of breast mortality rate, why CAD systems in breast cancer research are important in controlling breast mortality, different medical imaging modalities, possible abnormalities detected in mammograms, breast density classification schemes, and a description of microcalcification in mammograms. In this chapter, we discuss in detail the various techniques used for breast density estimation and classification, a detailed study of microcalcification classification methods, and the most common features used for microcalcification detection and classification. Likewise, review of texture analysis methods, and various texture feature descriptors are discussed. Additionally, a brief description of the databases used is given.

## 2.2 CAD in Medical Image Analysis - Mammography

Mammography is considered a gold standard method for early detection of breast cancer. Worldwide screening programmes have helped in reducing breast cancer mortality rates considerably in the last two decades. However, it is difficult to analyze a large number of mammographic images to detect a very small number of true positive cases due to the complex structure of breast tissue and parenchymal patterns along with dense tissue. Again there are chances of missing some cases due to the overlapping tissue structure thus hiding real abnormalities. The main errors during the radiological diagnosis can be divided into three categories, a) search errors (unnoticeable by the radiologist), b) detection errors, c) interpretation errors (lesion is detected but not classified correctly) (Calas et al., 2012). In addition to this, there has been a drastic

increase in the number of mammograms to be analyzed and this has posed a serious challenge to radiologists causing more diagnostic and classification problems along with inter- and intra- variability among radiologists in reading mammograms (Beam et al., 1996; Elmore et al., 1994). CAD (Computer-Aided Diagnosis) tools can be used to address some of the limitations in mammograms and reduce the variability between radiologists. Many clinical studies have shown that CAD systems have increased the sensitivity in breast cancer detection of about 20-21% compared to the radiologist (Baker et al., 2004; Brem et al., 2003; Destounis et al., 2009; Romero et al., 2011). Computer-Aided Detection (CADe) assists radiologists in finding cancer during screening and Computer-Aided Diagnosis (CADx) helps radiologists in deciding on the malignancy of abnormalities (Nishikawa, 2007).

A large number of CAD systems have been developed for assisting radiologists in screening mammogram (Doi, 2007; Nishikawa, 2007). In 1967, Winsberg et al. (1967) introduced a CADx system that determines a lesion in the mammogram as benign or malignant. Thereafter, a lot of studies were done to improve the performance of CAD systems to assists radiologists/doctors (Chan et al., 1990; Getty et al., 2000). The objective of CADe systems and CADe algorithms is not a diagnosis but to provide support to radiologists for improving the positive prediction from mammograms (Destounis et al., 2004; Getty et al., 2000; Horsch et al., 2006; Huo et al., 2002). Destounis et al. (2004) have analyzed the role of CADe in reducing false-negative results in mammogram screening through double reading. The results found that the CAD system was able to find 71% of the 52 findings diagnosed during screening which showed the potential of a CAD system in the screening process (Destounis et al., 2004). Various studies on CAD were performed on different parameters like lesion size (Brem et al., 2005b; Taylor et al., 2003), breast density, histological tumor type, calcifications, detection, recall, biopsy recommendation to see the potential of CAD in the screening process. Breast density was found to be one of the major problems affecting the sensitivity and performance of CAD system (Ho and Lam, 2003; Obenauer et al., 2006). High breast density can decrease the sensitivity of diagnosis, especially with the diagnosis of breast masses (Brem et al., 2005a; Ho and Lam, 2003; Obenauer et al., 2006; Sajeev et al., 2018). While detecting different tumor types Hadjiiski et al. (2006), it was found that the number of unnecessary biopsies by the radiologists was reduced by 0.7%. Similarly, different studies have shown that CAD had helped radiologists in detecting microcalcifications with a sensitivity ranging from 80% - 100% (Kallergi, 2004; Leichter et al., 2004; Soo et al., 2005).

Though the aim of CAD systems are not to replace radiologist for detection and classification, various studies were performed to analyze the role of CAD systems as a

double reader and single reader in screening mammograms. A study by Gilbert et al. (2006) in the UK National Breast Screening Programme, found that single reading with CAD systems in cancer detection was higher than through double reading. Around 6.5% more cancers were detected through single reading with a CAD system which in-turn increased the recall rate significantly from 15% to 32% from double reading. Similar studies showed the same suggestion that CAD systems help in improved cancer detection with increased sensitivity and (small) increase in recall rate (Ciatto et al., 2003; Gromet, 2008) for a large number of mammograms.

Recent studies showed that incorporating CAD systems in diagnosis can improve sensitivity and can enhance early-stage detection of malignancy (Castellino, 2005) though there is a debate that high sensitivity does not provide high detection rate by CAD systems (Lehman et al., 2015). It is estimated that the use of CAD systems in interpreting the digital screening of mammography has increased from 5% to 83% from 2003 to 2012 (Lehman et al., 2015) showing the advancement of CAD systems in medical research in diagnosis. Similarly, studies by Morra et al. (2015) showed that in digital breast tomosynthesis, CAD systems showed a larger detection rate of 89% for breast masses and microcalcifications with an acceptable FPR (2%) while the microcalcification cluster detection obtained a higher rate of sensitivity (95% with 37 of 39 microcalcifications detected).

The studies by Calas et al. (2012) showed that CAD systems are useful in cases where there is high inter-observer variability, lack of trained observers, difficulty in double reading by multiple radiologists. It is estimated that CAD performs better in detecting lesions of size between 1 and 3 cm (Brem et al., 2005b; Taylor et al., 2003). The studies showed that CAD is advantageous in dealing with patients with dense breasts because the sensitivity will decrease with the increase in breast density (Brem et al., 2005a; Ho and Lam, 2003; Obenauer et al., 2006).

While considering the detection and recall rates, Morton et al. (2006) observed an increase in the detection rate of 7.62%, Freer and Ulissey (2001) reported an increase of 19.5% of detection rate. The study by Warren Burhenne et al. (2000) analyzed 1,083 mammograms with biopsy-proven results and found that CAD system correctly detected 77% of the cases with no significant increase in recall rate (8.3% before and 7.6% after). When considering the biopsy recommendations, the decrease in the number of recommended biopsies by the use of CAD systems with an increase of accuracy of 2% shows the potential benefit of CAD systems in medical systems (Marx et al., 2004). Similarly, CAD systems are beneficial in cases where the interobserver variability is high. The studies by Jiang et al. (2001) showed that the variation in radiologist accuracy (interpreting clustered microcalcifications) was reduced to 46%

with the incorporation of CAD systems. Besides, the incorporation of CAD systems has decreased the average analysis time by the radiologist (Tchou et al., 2010). Thus the studies based on detection and recall rate, lesion size, processing time, breast density, sensitivity, classification accuracy, and recommended biopsies show that developing CAD systems are potentially useful in early detection and control of breast cancer mortality rate.

Overall the studies showed that CAD systems can play an important role in the mammogram screening process and have played a vital role in improving detection rate for large screening programmes and assisted radiologists in pointing out risk regions in mammograms for further investigation and reading. Similarly, advanced algorithms in CAD systems helped radiologists in image enhancement and detection for improving the specificity of cancer detection. Though CAD systems cannot be an alternative for an expert radiologist, it could help with increased sensitivity with a slight increase in recall rate and there is much more opportunity in improving CAD systems and CAD algorithms especially when dealing with dense mammograms.

## 2.3 Breast Density Classification and Risk Scoring

It is stated that risk in developing breast cancer in women ranges from many factors like hormones, genetics, age of first pregnancy, hormone replacement therapy (HRT), breast density, obesity, smoking and drinking, menopause, use of drugs, and age (Ferlay et al., 2013; Parkin and Fernández, 2006). Among these, breast density (Boyd et al., 2009) is considered as one of the top risk factors in developing breast cancer (Boyd et al., 2009, 2007; Maskarinec et al., 2005; Reporting, 2003). Breast density is the calculation of the relative amount of fibro-glandular tissue versus fatty tissue in the breast. Therefore, women with a higher percentage of fibrous tissue tend to have dense breasts with less fat tissue (Kumar et al., 2015). A high dense breast can hide the abnormality in a mammogram as it is a two-dimensional projection of breast tissue making a misperception among radiologists in diagnosing the abnormality accurately. Similarly, there are chances that various types of tissues overlap each other making it difficult to diagnose correctly (Ferlay et al., 2013; He et al., 2015; Virmani and Kumar, 2010; Wolfe, 1976b). Therefore, an intelligent CAD system should be able to increase the diagnostic accuracy of atypical cases masked by dense breast tissue (Kumar et al., 2015).

According to the literature, studies on breast density estimation and classification were started by Wolfe (1976b) based on the nature of breast parenchymal patterns. It is

observable that most CAD systems either follow an ROI based or segmentation based classification. It is stated that compared to global image measurement of breast density, the parenchymal texture descriptors based on localized descriptors can provide more refined features to measure the complexity and classification of breast parenchymal patterns (Gastounioti et al., 2016). Besides, parenchymal textural features could be considered as imaging markers to identify parenchymal changes concerning cancer development stages (Tan et al., 2013; Wang et al., 2010) and also can be used to grade breast malignancy (Bae et al., 2015; Sala et al., 2000).

It is to be noted that the majority of the methods in the literature for density segmentation are based on thresholding models, clustering models, or statistical models (He et al., 2015). He et al. (2015) have provided a detailed study on breast parenchyma segmentation methods based on these techniques. Torrent et al. (2008) had done a comparative study on two clustering-based algorithms and a region-based method for classifying dense and fatty mammograms. While the first strategy was based on multiple thresholding methods depending on excess entropy (Feldman and Crutchfield, 2003), the second approach used Fuzzy C-means clustering algorithm (Bezdek, 1981), and third focused on Fisherfaces method (FF) (Belhumeur et al., 1997) based on a statistical analysis of breast. The study provided a ROC analysis and concluded that clustering algorithms obtained better sensitivity.

Automatic single-threshold value methods can be used to separate highly dense tissue and low-intensity fatty tissue, but it fails due to mammographic density inhomogeneity. This lead to the proposal of adaptive thresholding methods for breast density segmentation (Kim et al., 2010; Neyhart et al., 2002; Zhou et al., 2001). Matsubara et al. (2001) proposed a model in which mammograms were divided into three regions using histogram analysis and discriminant analysis and classified into four categories based on the ratio of three regions. Similarly, Saha et al. (2001) proposed a model to classify digitized mammograms into dense/ fatty based using a scale-based fuzzy connectivity method. Features from image segments were extracted to characterize the mammographic density. To select the threshold, the minimization of energy function was computed based on the spatial arrangement of pixel intensities within the segmented region and across the regions. Sivaramakrishna et al. (2001) segmented the whole breast using a modified Kittler's optimal threshold procedure to segment dense regions. Zhou et al. (2001) classified breast density into four categories using a rule-based classification method by first segmenting the breast region using boundary tracking method and then an adaptive dynamic range compression technique was used to reduce the large grey level pixel variation to increase the differences in the characteristic features of the grey level histogram. Bovis and Singh (2002) proposed a new

method for the classification of mammographic images according to the breast type. The approach utilized texture features for classification by using Fourier transforms, laws texture masks, spatial grey level dependence (SGLD) using the directions {0°, 45°, 90°, 135°}, and discrete wavelet transforms as feature extractors. The Principal Component Analysis (PCA) was used for dimensionality reduction. To improve the classification results, a variation on bootstrap aggregation ('bagging') where multiple classifiers comprising feed-forward Artificial Neural Network (ANN) were trained using a ten-fold cross-validation scheme and using the combination rules proposed by Kittler et al. (1998). The method successfully classified mammographic images into BI-RADS four classes obtaining an accuracy of 71.4%.

Masek et al. (2003) used average histograms of mammographic images of each density class in MIAS dataset for classification and accuracy of 62.42% was obtained using Euclidean distance measure. Similarly, Zwiggelaar et al. (2005) used statistical grey-level histogram modeling (PCA) for triple MIAS density classification and obtained an accuracy of 71.5%. Later, Muhimmah and Zwiggelaar (2006) proposed quantitative estimation approaches based on histogram information to classify the mammographic images. A multi-resolution histogram technique that used Directed Acyclic Graph-Support Vector Machine (DAG-SVM) classifier to utilize texture feature was proposed. The approach successfully classified the MIAS dataset obtaining an accuracy of 77.5%.

As discussed, density classification using the fibroglandular disc region in mammogram attained a vital role in classification. Ferrari et al. (2004) proposed a method to segment the fibro-glandular disc in mammograms based on a statistical model of breast density using a mixture of up to four weighted Gaussians. The parameters of the model and number of tissue classes in the breast were estimated using an expectation-maximization algorithm (EM) and minimum description length method along with grey level statistics of the pectoral muscle to determine the tissue categories. Later, El-Zaart et al. (El-Zaart, 2010) developed a similar method using Gamma distribution instead of Gaussian distribution for modeling data in mammography images. The histograms of images were seen as a mixture of Gamma distributions and the expectation-maximization technique with a gamma distribution (EMTG) was proposed to estimate the statistical histogram parameters. It was found that the Gamma distribution model performed better than the Gaussian mixture model. Torres-Mejía et al. (2005) performed a comparison of qualitative and quantitative evaluation using data from the UK Guernsey. The mammograms were classified using subjective qualitative Wolfe classification and many quantitative mammographic features using computer-based techniques. The study showed that breast cancer risk is affected not only by the

amount of mammographic but also by the degree of heterogeneity of the parenchymal pattern.

Karssemeijer (1998) proposed a method based on the distance transform that subdivided the breast tissue area into regions for automated determination of parenchymal patterns in mammograms. The study was performed to study the relation between breast cancer risk and changes in mammographic density. The grey level histogram features were computed from the subdivided regions along with other features like the differences between tissue projected in pectoral and breast area and the KNN classifier was used for classification. Bosch et al. (2006) proposed an approach to model and classify breast parenchymal tissue. The distribution of tissue densities was investigated using different descriptors like texture and SIFT features, where textons performed better than SIFT. The features were classified using a classifier based on local descriptors and probabilistic Latent Semantic Analysis (pLSA). Martin et al. (2006) developed a method (computer-aided mammographic density estimation-MDEST) for density measurement using the percentage density where pixel grey-levels in breast borders were analyzed and dense areas were segmented. To calculate the percentage density, the number of dense pixels was divided by the total number of pixels within the borders. The results found that the MDEST has cross-correlation with the radiologist-estimated density, although there was a large overlap between percentage density in qualitative BI-RADS categories 2-4.

Miller and Astley (1992) investigated texture-based discrimination between glandular and fatty regions by applying granulometric techniques and Laws texture masks. Similarly, Petroudi et al. (2003) used maximum-response filter bank (MR8), which mapped to a texton histogram to capture the mammographic appearance to perform breast parenchymal pattern classification. Oliver et al. (2005b) proposed a method to classify mammogram images into BI-RADS density classes based on their breast parenchymal density, which extracts the breast tissue texture. This approach groups the pixels with similar tissue having similar grey-level values using fuzzy c- means algorithm (Bezdek, 1981). Morphological features like relative area, the center of masses, and medium intensity from both clusters and texture features from co-occurrence matrices were used for classification using KNN (Duda et al., 2001). Similarly, the same approach was used to classify the MIAS triple class using KNN and ID3 classifiers (Oliver et al., 2005a). Likewise, Oliver et al. (2006) performed a comparative study between different methods like a fractal, statistical and fuzzy-c means and showed that the statistical approach obtained more significant and separated regions when fractal approach resulted in a pixelated segmentation. It was noted that small tissue variations were not affecting the density classes assigned by statistical method, but it

can affect the results of the fuzzy c-means algorithm. Oliver et al. (2007) later used clustering methods like Fuzzy-c-means, Normalized cuts algorithm (Shi and Malik, 2000), and Mean Shift algorithm (Comaniciu and Meer, 2002) to cluster/segment the tissue regions. The texture descriptors like Local Binary Patterns and Co-occurrence Matrices were used for texture feature extraction. Normalized Cuts performed lower when compared to others due to the inherent nature of the algorithm, while the LBP and co-occurrence matrices performed almost in a similar manner.

Later, Oliver et al. (2008) developed an automatic breast tissue classification methodology based on segmentation of breast area into dense and fatty tissue, extraction of textural and morphological features followed by classification scheme by using a Bayesian combination of different classifiers. The method showed a high correlation between automatic and expert-based BI-RADS density assessment. The same approach was used later in full-field digital mammograms (FFDM) and obtained similar results (Tortajada et al., 2012).

Subashini et al. (2010) proposed a method where ROI was restricted to breast tissue. The technique included preprocessing, statistical feature extraction, and classification through the SVM classifier. Grey level thresholding and connected component labeling were used to estimate the artifacts and pectoral muscle from ROI. The method obtained an accuracy of 95.44% for a MIAS triple class density classification.

Blot and Zwiggelaar (2001) performed classification of mammographic parenchymal patterns where the background texture and structures in images were separated. The approach was based on the statistical difference between local and median co-occurrence matrices. The method focused on background texture information. The study showed that although there was no improvement in classification results compared to Karssemeijer's technique (Karssemeijer, 1998), there was no loss of information during the reconstruction of texture. Liu et al. (2010) used histograms of multiple resolution mammograms to classify density using the SVM classifier. Similarly, statistical features, morphological and texture features were used to classify mammograms based on breast tissue/density (Chatzistergos et al., 2010; He et al., 2012; Kim and Kim, 2014; Papaevangelou et al., 2011; Sharma, 2017; Vállez et al., 2011).

Oliver et al. (2015) proposed an automated approach for breast density segmentation in mammographic images. The method used a supervised pixel-based classification and textural and morphological features for density segmentation. The study showed that the dense tissue percentage decreased over time. Chen and Zwiggelaar (2010) developed a breast region segmentation method using histogram thresholding, active contour model, and region growing, which was later used for mammographic density estimation. The density estimation was done using a modified fuzzy-c means

method where spatial information was incorporated into traditional FCM clustering algorithmp Chen et al. (2011). Similarly, a topographic model where a topographic map was constructed to represent the regions having a range of densities (Chen et al., 2011). Additionally, Chen et al. (2013b) developed a multiscale blob based model for representing mammographic patterns. Ben-Ari et al. (2016) proposed a breast segmentation method based on training with weakly labeled data (breast density categories). The method introduces a segmentation strategy based on the correlation between segmentation parameters and breast density. The threshold for segmentation is estimated using a recursive approach based on the Fuzzy-Logic framework. The method obtained an accuracy of 75% using 243 full-field digital mammograms.

Antonelli et al. (2019) proposed descriptors based on Permutation Entropy (PE) algorithms for mammographic density estimation. When the characteristic vector for each mammogram is obtained, they were trained using a multi-layer neural network. Ahn et al. (2017) classified breast density using a Convolutional Neural Network (CNN) using the 397 full-field digital mammogram database provided by Seoul National University Hospital.

Virmani et al. (2016) used features from Law's texture energy from Laws' masks of length 5, 7, and 9 for classifying MIAS mammograms into fatty and dense tissue. The dimensionality reduction was performed using PCA, classification using SVM and Probabilistic neural network (PNN), and attained an accuracy of 94.4% with the first four principal components derived from features for Law's mask of length five.

Cai et al. (2018b) performed a breast density estimation using a wavelet transform. The grey-scale histogram of the breast with different densities is enhanced by adaptive histogram equalization through wavelet fusion methods. Then the threshold of the mammary region was computed using a grey-scale histogram, which later was classified into four categories. The method obtained an accuracy of 85% using 400 mammograms from a private dataset. Fowler et al. (2014) developed an automated percentage of breast density measurement (PDa), which analyzes signal-dependent noise (SDN) based on a wavelet expansion. The application used Full Field digital mammography (FFDM) for performance evaluation.

Wei et al. (2018) proposed a supervised deep learning approach for automated estimation of percentage density (PD) in digital mammography. The Deep Convolutional Neural Network (DCNN) was trained to estimate a probability map of breast density (PMD) where the PD was computed as the ratio of dense area to the breast area. The method was successful in finding a strong correlation with DCNN estimation and segmentation by radiologists. Kallenberg et al. (2016) proposed an unsupervised deep learning method for density estimation and risk scoring. The proposed convo-

lutional sparse autoencoder (CSAE) model consisted of a sparse autoencoder within a convolutional architecture. Mohamed et al. (2018) constructed a convolutional neural network (CNN)-based model to classify digital mammograms into BI-RADS "scattered density" and "heterogeneously dense" categories. The approach obtained an $A_z$ of 0.9421.

Similar to BI-RADS and Wolfe density classification schemes, a few studies classified mammographic images into Tabar classification (He and Zwiggelaar, 2013; He et al., 2008b, 2009, 2010a; Muhimmah et al., 2007). He et al. (2009) performed mammographic segmentation based on spatial moments and prior information on mammographic building blocks like nodular, linear, homogenous, and radiolucent.

It is observed that most of the breast density segmentation/estimation methods utilized the texture features, statistical and morphological characteristics of breast tissue. Though there are a few deep learning approaches in the literature, most of them focus on FFDM images and classify the density into two categories, unlike the BI-RADS four-category or MIAS triple category. Similarly, some studies show that image enhancement approaches can improve the density estimation as it gives more information on the underlying tissue structure. Few other methods based on LBP and breast density modeling are explained in respective chapters 3 and 4.

## 2.4 Classification of Microcalcification

### 2.4.1 Microcalcification- BI-RADS Classification Descriptors

The presence of microcalcification clusters is considered as a primary sign in breast cancer detection (Sakka et al., 2006). Microcalcifications are deemed to be one of the prominent symptoms in indicating the carcinomas (Kallergi, 2004). 55% of non-palpable cancers are diagnosed by the presence of mammographic microcalcifications (Gajdos et al., 2002; Hernández et al., 2016; Holland and Hendriks, 1994). Though breast microcalcifications are common findings in mammography, the morphology of individual micro-calcifications (MC), MC-cluster and distribution of microcalcifications are considered to be the most critical features noticed by radiologists for the classification of microcalcifications as benign or malignant (Elter and Horsch, 2009; Hernández et al., 2016). The primary descriptors as per BI-RADS categories of microcalcifications based on $5^{th}$ version showed the prominence of morphology and distribution in classifying microcalcifications (Sickles et al., 2013).

BI-RADS microcalcification distribution descriptors defined the arrangement of calcifications inside the breast region (Hernández et al., 2016). The distribution de-

scriptors can be diffuse distribution (scattered within the breast), regional distribution (pattern in an extensive area of above 2cm dimension), grouped/clustered distribution (minimum of 5 calcifications in 1 cm to more calcifications within 2cm), linear distribution (calcifications are arranged in the linear pattern) and segmental distribution (calcium deposits in the ducts and its branches). BI-RADS microcalcification morphological descriptors explain the morphology of calcifications within the breast tissue.

The benign calcifications can be vascular (usually calcium deposits in the walls of mammary arteries forming a bilateral railroad shape), skin/dermal calcifications (calcifications on the sebaceous gland), milk of calcium calcifications (small particles of calcium oxalate deposits in saccular dilatations of TDLU), extensive linear calcifications (secretory or road like calcifications located in ducts), popcorn calcifications (dense, thick and more massive), dystrophic calcifications (usually due to the existence of fat), round/punctate calcifications, ring calcifications (egg shelled or radiolucent centered) and suture calcifications (calcium deposited in suture material, with linear or tubular calcifications with knots).

The morphological descriptors for suspicious calcifications include coarse heterogeneous calcifications (rough, irregular, varied and may be located in breast ducts), amorphous calcifications (powdery, cloud or cottony pattern with less than 0.1 mm size), fine pleomorphic calcifications (called as crushed stone with different size, shape, and heterogeneous nature) and fine linear or branched calcifications (small-sized with less than 0.5 mm, thin, linear and irregular edges in a discontinuous pattern. They tend to branch in different directions forming shapes of letters like L, V, Y, X). A detailed classification according to BI-RADS categories ($5^{th}$ version) is summarized in Table 2.1.

### 2.4.2 Microcalcification Detection/Classification

As discussed, micro-calcifications appear to be bright spots in mammograms. Depending on the morphology and distributional pattern of micro-calcifications as discussed (see Section 2.4.1), they can be classified as malignant or benign cases (Hernández et al., 2016). According to the studies, the presence of microcalcification clusters is considered to be a vital sign for the early detection of breast cancer. So, the high correlation between microcalcification clusters and cancer shows the importance of developing CAD systems for automatic classification of micro-calcifications (Cheng et al., 2003; Sakka et al., 2006). MCCs are a common abnormality found in mammograms and are found in 30-50% of mammograms during screening, which could

Table 2.1 Classification of calcifications according to BI-RADS categories (Hernández et al., 2016).

| Calcification type | Category |
|---|---|
| Vascular calcifications | |
| Skin calcifications | |
| Thick linear calcifications | |
| Popcorn calcifications | BI-RADS 2 |
| Dystrophic calcifications | |
| Round, scattered or isolated calcifications | |
| Ring calcifications | |
| Suture calcifications | |
| Round grouped calcifications | BI-RADS 3 |
| Coarse, rough, heterogeneous calcifications | BI-RADS 4 B |
| Amorphous calcifications | BI-RADS 4 B |
| Fine pleomorphic calcifications | BI-RADS 4 B |
| Linear or branched linear calcifications | BI-RADS 4 C |
| Linear and new branching linear and segmental distribution calcifications | BI-RADS 5 |

promise up to 60-80% cancer cases during histological examination (Cheng et al., 2003; Ren, 2012; Sentelle et al., 2002). A typical CAD system includes preprocessing, ROI extraction, detection/segmentation of microcalcifications, and classification of microcalcifications as benign/ malignant (Ren, 2012).

Preprocessing or enhancement techniques could be used to enhance the intensity/contrast, remove background noise, and sharpen the ROI edges or boundaries (Beghdadi and Le Negrate, 1989; Cheng et al., 2003). The image enhancement techniques could be based on conventional methods like masking using global or fixed-sized windows by amplifying the high-frequency pixels and reducing low-frequency data (Narod et al., 2015; Papadopoulos et al., 2008; Singh and Kaur, 2017). The region-based enhancement methods where a pixel is taken as a seed point and the region is enhanced and grown, based on the pixels, region's contrast and background (Mohideen and Thangavel, 2013; Morrow et al., 1992). The feature-based enhancement methods using fractals (Li et al., 1997; Stojić and Reljin, 2010) and wavelets (Mustra et al., 2012) for preserving the sharpness and shape of the microcalcification spots are the typically used enhancement approaches.

The second stage of CAD system involves the ROI extraction and classification, which could be done to assist radiologists with suspicious regions or to classify the abnormalities detected as benign or malignant (Cheng et al., 2003; Li et al., 1997). The feature/primitive extraction of abnormality is a vital step in the detection and classification of abnormality into benign or malignant causes. The main features extracted to categorize microcalcifications are based on the morphology, texture, and cluster

features. According to the BI-RADS classification category, the abnormality type could be benign (B), malignant (M), or Uncertain (U) (Kallergi, 2004). Kallergi (2004) and Muttarak et al. (2009) explained a detailed classification based on morphology, distribution and the number of MCs. The microcalcifications appear as clusters in mammogram (Karssemeijer, 1991; Veldkamp et al., 2000). Therefore, radiologist utilizes the characteristics cluster features of microcalcifications like morphology of individual and cluster, location of the cluster, and distribution for reporting the malignancy of the abnormality (Cheng et al., 2003; Elter and Horsch, 2009).

Individual microcalcification feature extraction included properties like perimeter, compactness, area, elongation, eccentricity, contrast, orientation methods (Cheng et al., 2003; Ciecholewski, 2017; Roty et al., 2017; Yu and Guan, 2000), grey level features, morphological features (Betal et al., 1997; Ciecholewski, 2017; Papavasileiou et al., 2018; Suhail et al., 2018a; Woods et al., 1993), using multiple threshold levels (Dominguez and Nandi, 2007) and estimating the probability of each pixel type to be a part of lesion using Gaussian mixture models (Singh and Bovis, 2006). Karssemeijer (1991) developed statistical methods and Bayesian image analysis for extracting clustered microcalcifications from mammograms, Yu and Guan (2000) used mixed features from wavelet features and grey level statistical features based on connectivity for detection and classification, Chan et al. (1990) used the local and global thresholding methods for detection and classification. In order to study the tissue around the microcalcification in classifying the microcalcifications, Karahaliou et al. (2008) extracted texture features and attained an $A_z$ of 0.98. Texture features of microcalcification tissue and surrounding tissue were investigated by researchers for classification of microcalcification using feature extraction methods like co-occurrence features (DaPonte and Sherman, 1991; Dhawan et al., 1996; Enderwick and Micheli-Tzanakou, 1997; Karahaliou et al., 2008; Yu and Guan, 2000; Zyout and Abdel-Qader, 2011), grey level run length features (Kim and Park, 1999; Kramer and Aghdasi, 1999), Gabor filter bank (Bhangale et al., 2000; Farzana and Rathnakara, 2011; Rogova et al., 1999), spatial grey level dependence (SGLD) matrices (Geronimo et al., 1994; Hamdi et al., 2008; Kim and Park, 1999), DT CWT (Tirtajaya and Santika, 2010), Laws features (Dheeba and Selvi, 2011; Karahaliou et al., 2008; Laws, 1980b) feature extractors and obtained good classification results. Similarly, morphological shape and related properties of individual MCCs and cluster were investigated for MC classification (Betal et al., 1997; D'Aloia et al., 2013; Soltanian-Zadeh et al., 2004) using statistical (Soltanian-Zadeh et al., 2004), Fourier descriptors (Chitre et al., 1993; Ma et al., 2010; Shen et al., 1993, 1994), Radon transform (Bocchi and Nori, 2007),

wavelet features like entropy, energy (Kocur et al., 1996; Malar et al., 2012; Sakka et al., 2006; Soltanian-Zadeh et al., 2004; Tirtajaya and Santika, 2010).

Other than the morphology of microcalcification, distribution and number of microcalcification are vital features in classification of MCCs (Chen et al., 2015; Cheng et al., 2003; Elter and Horsch, 2009; Feig et al., 1987; Kallergi, 2004; Ren, 2012; Sickles, 1986; Suhail et al., 2018b). The proximity of microcalcification distribution is an important indication of the type of abnormality of MCCs. It has been reported that the malignant MCCs tend to be smaller in size, densely distributed occupying more than five microcalcifications within $1cm^2$. In contrast, benign MCCs tend to be larger and widely spread (<= 5 calcifications per 1 $cm^2$) (Chen et al., 2015; Cheng et al., 2003; Feig et al., 1987; Ren, 2012; Sickles, 1986; Strange et al., 2014; Suhail et al., 2018b). Therefore, investigating cluster features like location of a cluster, distribution of individual microcalcifications and cluster morphology has been an important factor in MCCs classification. While considering the location of microcalcification cluster, it is a well-established fact that malignant lesions are widely seen in the upper outer quadrant of the breast (ACR, 2003a,b). Veldkamp et al. (2000) automatically detected the microcalcifications using statistical methods based on Bayesian and Markov random field. Later, the relative distance of cluster to pectoral muscle and breast edge was estimated to classify the clusters and obtained an $A_z$ of 0.83 for a subset of 90 images. Similarly, Russakoff and Hasegawa (2006) investigated the location of masses in the breast region and revealed areas of higher risk of breast cancer, though the location of MCs were not considered. Later, Andreadis et al. (2015) investigated the location of MC cluster to develop a probabilistic map to determine the risk of malignancy. The main disadvantage in determining the malignancy using cluster location is that there are limitations with the images under investigation as it needs prerequisites like information on the location of the breast skin, pectoral muscle and nipple. Later, Ashiru and Zwiggelaar (2016) conducted a study based on classification using location (Andreadis et al., 2015) and topological modelling based on Chen et al. (2015) and found topological modelling performed better than location-based classification.

While considering the distribution of individual classification in a cluster region, the spatial distribution through mean and standard deviations were employed along with eccentricity and normalized central moments of the MC centroids (Alam et al., 2018; Betal et al., 1997; Dhawan et al., 1996; Lo et al., 2003; Papadopoulos et al., 2005; Schmidt et al., 1999; Soltanian-Zadeh et al., 2001, 2004; Suhail et al., 2018a,b; Wei et al., 2005a). Similarly, several approaches have been developed for discriminating the MCCs based on surrounding tissue/background tissue assuming that the abnormality

could change tissue structure of breast (Elter and Horsch, 2009). Texture features from surrounding tissue region of the cluster were analysed through grey-level co-occurrence matrices (Dhawan et al., 1996; Karahaliou et al., 2007; Kramer and Aghdasi, 1999; Pratiwi et al., 2015; Sakka et al., 2006), Haralick features (Fu et al., 2005; Geetha et al., 2008; Lucas and Mortezaie, 2002; Thiele et al., 1996), wavelets and multiwavelet features (Kramer and Aghdasi, 1998; Soltanian-Zadeh et al., 2004).

Similarly, topological modelling of microcalcification based on spatial distribution was investigated (Alam et al., 2018; Bocchi et al., 2004; Chen et al., 2014, 2015; George et al., 2018a; Rampun et al., 2018b; Strange et al., 2014; Suhail et al., 2018a,b). Chen et al. (2014, 2015) investigated the topology and connectivity of microcalcification in clusters for classifying the abnormality into benign or malignant. They performed morphological dilation operation on microcalcifications at multiple scales to create a subgraph based on overlapping of micro-calcifications during the dilation process. After that, eight graph features were extracted to classify the abnormality. Similarly Alam et al. (2018), extracted graph connectivity features at particular scales based on the denseness of cluster at different fixed scales of graph topology. Likewise, Strange et al. (2014) used topological modelling and introduced mereotopological barcode for classifying microcalcifications into benign or malignant. They performed a morphological operation (RCC8D, 8 regions Connected Calculus) on segmented microcalcifications at multiple scales to investigate the connectivity and developed a barcode for classifying microcalcifications into benign or malignant. Later, Suhail et al. (2018b) developed a topological model based on binary tree structure properties and classified microcalcification clusters as benign or malignant based on the height of the tree. The scope of topological models was investigated by Ashiru and Zwiggelaar (2016) and found that topological model performed better than location-based, which gives more closeness from a clinical perspective of evaluation.

All the above methods based on topology were performed on segmented micro-calcifications from DDSM and MIAS datasets which helped us in providing a direct comparison with the discussed methods. The thesis introduces a new topological model based on the distribution of microcalcifications in an ROI called connected-chain model. The connected chain at multiple scales is evaluated to make the method resolution-independent. The distributional closeness of microcalcifications was estimated using mathematical morphological operations. Similarly, we have investigated the surrounding tissue features based on various texture feature methods. Then connected-chain topological model was compared to analyse the potential of topological models in the CAD systems for micro-calcification classification.

A wide variety of classification methods are used to classify micro-calcifications into benign or malignant using the extracted features. A wide range of machine learning algorithms helps the CAD systems in discriminating the malignant and benign microcalcifications from the extracted features. The most common machine learning algorithms used by researchers are algorithms based on Neural Networks, K-nearest neighbour classifier, Bayesian-based classifiers, linear classifiers, fuzzy decision tree classifiers, Binary decision tree classifiers, Support Vector Machines. Kramer and Aghdasi (1998) used KNN classifier to distinguish the abnormality using the features extracted through co-occurrence matrices and wavelet features. Later, a comparative study by Kramer and Aghdasi (1999) compared the performances of KNN and Neural Network (NN) on the texture features extracted from Nijmegen dataset and found that NN performs better than KNN. Similarly Soltanian-Zadeh et al. (2001) used KNN to classify a set of shape feature using Nijmegen dataset to discriminate the abnormality and obtained an $A_z$ of 0.82. Bhangale et al. (2000) used a KNN classifier to classify the texture features extracted from microcalcifications using Gabor filter banks.

Similar classifications were done using KNN classifier by various other studies for microcalcification classification (Chen et al., 2015; Engan et al., 2007; Rad et al., 1999). Yilmaz et al. (2002) proposed an approach based on Support Vector Machine (SVM) for detection and classification of microcalcification clusters. Ren (2012) extracted cluster features to classify microcalcifications using SVM and ANN. They introduced a method to deal with unbalanced data which is a prevalent problem in medical image classification as the number of cases of a particular type of abnormality will be very high compared to others. Similarly, SVM was used by various researchers for classification of benign or malignant microcalcifications (Dheeba and Selvi, 2011; Fu et al., 2005; Singh et al., 2006; Wei et al., 2005a,b). The next common classifier under consideration was Bayesian classifier which was used (Karabatak, 2015; Krishnaveni et al., 2014; Patrocinio et al., 2004; Zyout et al., 2009) for investigating the malignancy of microcalcification cluster. Bankman et al. (1994) extracted features that represented individual microstructures using the contour map and used a neural network for the classification.

Likewise, the classification capability of the neural network was utilized for better classification of microcalcifications by various researchers (Chan et al., 1997; Khehra and Pharwaha, 2016; Ren, 2012; Shachor et al., 2018) and advanced models of neural networks were developed for better classification. Deep learning neural networks are advancing its architecture in better classification of features, and it finds its best application in micro-calcification classification. Samala et al. (2016) used deep learning CNN for classifying microcalcification in tomosynthesis, Bekker et al. (2016) used a

multi-view deep learning architecture for microcalcification classification, Wang et al. (2016) used deep learning for discriminating both masses and microcalcifications, Cai et al. (2018a) proposed a fully automatic microcalcification detection based on deep-CNN.

In this thesis, we will analyze the microcalcification classification using the most commonly used classifiers like KNN, Bayesian classifier, SVM for evaluating the performance of the proposed microcalcification feature extraction and classification methods. Thereby it provides a direct comparison with other topological modelling methods in the literature. A brief description of the standard features extracted and classifiers used for classifying the microcalcifications into benign or malignant cases are explained in Table 5.10, which shows a comparison with the proposed methods. Separate comparative tables are shown in the respective sections in chapter 5.

## 2.5 Texture

### What is Texture?

Texture cannot be defined mathematically, or with a comprehensive definition, it can be considered as a fuzzy concept as it depends on various human interpretations. According to Tuceryan and Jain (1993), a universal definition cannot be given to texture, but the definition depends on the area of application. Image texture is considered as a set of metrics calculated in image processing giving information about the arrangement of colour or intensities of a selected region of the image in a space (Linda and Shapiro, 2001). Sklansky (1978) defines the texture of a region as "A region in an image has a constant texture if a set of local statistics or other local properties of the picture function are constant, slowly varying, or approximately periodic." Similarly, Tamura et al. (1978) define texture from an analytical viewpoint as" We may regard texture as what constitutes a macroscopic region. Its structure is attributed to the repetitive patterns in which elements or primitives are arranged according to a placement rule (Rosenfeld, 1975)". According to Haralick (1979), image texture can be considered as an organized area phenomena which on decomposing may have two dimensions. While the first dimension describes the primitives out of which the image is composed, the second dimension explains spatial dependence or interaction between image primitives. It gave a structural approach to the definition of image texture. Similarly, Tamura et al. (1978) attempted to explore statistical features to texture attributes resulting in explaining coarseness, contrast, regularity, roughness, directionality and line-likeness properties of image texture.

Table 2.2 A brief list of microcalcification classification methods.

| Feature | Database | Cases | Classifier | Method | Result |
|---|---|---|---|---|---|
| Law features | MIAS | 322 | SVM | Dheeba and Selvi (2011) | CA=86.1% |
| Intensity, shape and linear structures | DDSM | 150 | ANN/SVM | Ren (2012) | $A_z$=0.94 |
| Shape | DDSM | 183 | Threshold | Ma et al. (2010) | $A_z$=0.96 |
| Topological features | DDSM | 300 | kNN | Strange et al. (2014) | CA=80% |
| Cluster | MIAS | 25 | SVM | Papadopoulos et al. (2005) | $A_z$=0.81 |
| Texture around microcalcifications using grey level first order, GLCM, Laws | DDSM | 100 | Neural network and Majority Voting | Karahaliou et al. (2008) | $A_z$=0.98 |
| Texture, Shape using SGLD | MIAS | 22 | Fisherpsilas linear discriminant | Hamdi et al. (2008) | CA= 95.5% |
| Texture ,DT CWT | MIAS | 27 | SVM | Tirtajaya and Santika (2010) | CA=88.64% |
| Texture | MIAS | 25 | KNN | Zyout and Abdel-Qader (2011) | $A_z$=0.88 |
| Texture using Gabor | Private | 35 | PNN | Farzana and Rathnakara (2011) | $A_z$=89.06 |
| Location of cluster | Nijmegen | 90 | KNN | Veldkamp et al. (2000) | $A_z$=0.83 |
| Texture-co-occurrence features | DDSM | 85 | KNN | Karahaliou et al. (2007) | $A_z$=0.96 |
| Texture | unknown | 54 | Logistic discriminant analysis | Thiele et al. (1996) | Sensitivity 89%, Specificity 83% |
| Shape and texture | Nijmegen | 40 | KNN | Soltanian-Zadeh et al. (2001) | $A_z$=0.82 for shape, $A_z$=0.72 for texture |
| Topology & location | MIAS | 20 | kNN | Ashiru et al. Ashiru and Zwiggelaar (2016) | $A_z$=0.95 |
| Topology & location | DDSM | 280 | kNN | Ashiru and Zwiggelaar (2016) | $A_z$=0.75 |
| Tree-based modelling | DDSM | 129 | majority voting | Suhail et al. (2018b) | CA=91% |
| Topology | DDSM | 300 | kNN | Chen et al. (2015) | CA=85.2 % ± 5.7 % |

Based on spatial homogeneity, image textures can be classified as homogeneous (texture contains uniformly repetitive structures producing idealized patterns), weakly-homogeneous (explains local spatial variation in the spatial arrangement of texture elements violating the precise repetitiveness as homogeneous), and in-homogeneous/ heterogeneous (the self-similarity of repetition of texture elements are absent). Figure 2.1 shows sample images of texture based on homogeneity.

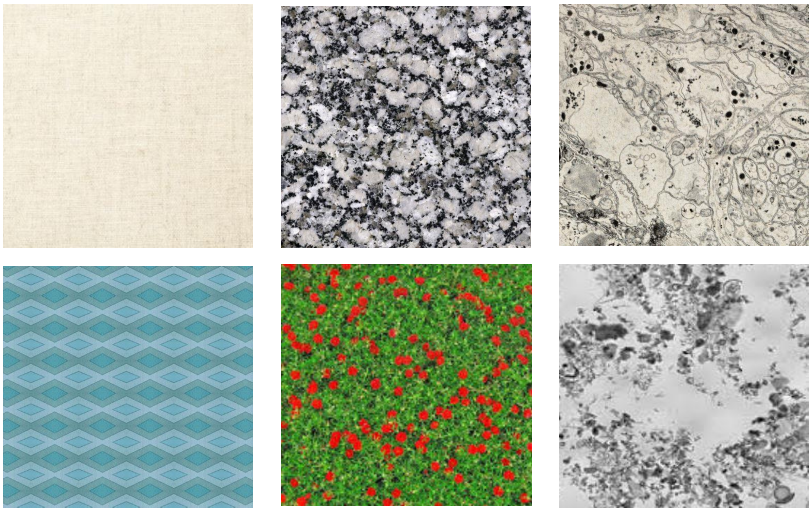


Fig. 2.1 Representation of texture based on homogeneity: The first column represents homogeneous texture, second column represents weakly-homogeneous texture and third column represents in-homogeneous texture images.

Texture analysis can be explained as the process of extracting meaningful information or details from the surface of an object or structure/patterns appearing in an image or video, where the object could be representing a small region or whole image.

## 2.6 Texture Analysis Techniques

The image processing procedure is divided into different stages. Noisy image regions are pre-processed, followed by feature extraction where texture patterns are extracted/segmented for classification. The efficiency of the texture analysis process can be calculated by the accuracy of classifying patterns or identifying objects based on extracted features. So, developing an effective feature extraction algorithm or a robust classifier can provide high-quality features and improve accuracy, respectively. Feature extraction techniques can vary depending on the application and the spatial relationship between the arrangement of pixels in an image region is observed (Tuceryan and Jain, 1993). Some of the feature extraction methods can include spatial features, transform

features, edges and boundaries, shape features, moments and texture features. Though there is great diversity in texture feature extraction techniques, we focus on the most widely used techniques of texture analysis in this section classifying them into four categories:

- Statistical methods: Statistical methods explain the distribution and relationship between grey-level pixel values.
- Structural methods: Structural methods explain texture features based on texture elements called texels/primitives and the placement rules of elements.
- Signal processing/filtering based approaches: Methods based on applying a group of filters to images in the frequency or spatial domain.
- Model-based techniques: Approaches define texture as a probabilistic model.

### 2.6.1 Statistical Methods

Statistical-based techniques explain the texture using properties governing the distribution and relationship between grey-level pixel values in the image region. According to Haralick (1979), an image texture is considered non-figurative and cellular. The texture is composed of two dimensions on decomposing, where the first dimension describes the primitives, which make the image texture while the second dimension represents the spatial representation of primitives. Depending on the number of image pixels defining the local feature, the approaches can be further classified into first-order (single-pixel), second-order (two pixels under consideration) or higher-order (more than three pixels) statistics (Ojala and Pietikäinen, 2004). While first-order approaches calculate texture properties like average and variance of pixels, higher-order calculates features between multiple pixels in a specific location relative to each other. A commonly used statistical description of texture is the intensity histogram estimation which is the first-order statistical analysis of the image. The most common texture features extracted from intensity histogram are mean, variance, skewness, kurtosis, energy and entropy, which provide the grey-level information of pixels in the image. However, the histogram of intensity levels of an image gives only a summary of the statistical information about an image or fragment of the image. The histogram shape can reveal many features, e.g. a narrowly distributed histogram represents the low-contrast of image. Though the histograms calculate the statistical properties of individual pixels, it cannot provide any details about the relative position of pixels to each other or fragments in an image.

Haralick et al. (1973) proposed a second-order statistical feature extraction method using co-occurrence matrix where the spatial relationship of grey-levels between a pair of pixels are analyzed at a particular distance or angular spatial distance. They were able to extract potential texture features for texture classification, especially in bio-medical images (Lerski et al., 1993; Pratiwi et al., 2015). A wide range of texture features can be extracted from this grey-level matrix to measure texture properties within the image region, such as homogeneity, contrast and correlation which made it famous and widely used texture extraction approach. A wide range of models based on the grey-level co-occurrence matrix was developed in the literature to extract the two-dimensional spatial relationship between pixels in an image. Lam (1996) developed grey-level gradient co-occurrence matrix (GLGCM) to capture the second-order statistics of grey level gradients between pixels. Similarly, other variants like grey level difference matrix (GLDM) estimated the absolute difference between grey levels at a predefined distance. GLDM can extract features similar to GLCM like contrast, angular second moment, entropy, mean and inverse difference (Weszka et al., 1976). Galloway (1974) proposed a new variant grey-level run-length matrix (GLRLM), where the number of grey level runs of specific length is computed. The grey level run represents the linearly adjacent pixels with the same pixel value and length denotes the number of pixels in a run. In order to evaluate the spatial distribution of pixels estimating the probability density function (PDF, estimate the grey level values of two pixels at a predefined distance and direction), Spatial Grey Level Dependency (SGLD) matrices were developed (Clausi, 2002; Galloway, 1974; Weszka et al., 1976).

A detailed description of statistical texture features will be explained in Section 2.8

### 2.6.2 Structural Methods

A common structural approach of texture analysis is fractal analysis (Kaplan, 1999; Wohlberg and De Jager, 1999; Xia et al., 2006; Xu et al., 2006). A fractal is a structure where similar patterns appear recursively at progressively smaller scales and maintain its shape at all scales (e.g. snowflakes). The most common method for calculating the fractal dimension is the box-counting method (Li et al., 2009; Sarkar and Chaudhuri, 1994). The basic approach of box-counting is, for an object $E$ in an image, the image is covered by boxes and the number of boxes occupied by object $E$ is counted. Lu and Fu (1978) developed a syntactic structural method for extracting and discriminating textures. The texture pattern is divided into fixed-size windows. The windows belonging to same pattern units are characterized by a tree grammar which is later used to synthesize and discriminate the textures. Similarly, Ahuja

(1982) developed a model to define the neighbourhood of a point known as Voronoi tessellation. The Voronoi polygons assumed to have appealing characteristics as expected from the neighbourhood of a point. The geometrical features of the Voronoi neighbourhoods are used as characteristic features in dot pattern processing which on analysis incorporates joint properties of neighbour points. Later, Tuceryan and Jain (1990) introduced a texture analysis based on Voronoi tessellation. They defined texture in terms of primitives called tokens. The algorithm created a Voronoi tessellation of the tokens to define the textured image. Then it computes feature vector for each of the Voronoi polygons to identify the interior and border regions of textures.

The main drawback of the structural approach is determining how to define texels that represent texture structures, especially when image structure does not follow a particular structure or pattern. So these approaches are better suited for a significant pattern (macro-structures) than micro-structure texture images.

### 2.6.3 Signal Processing/Filtering Based Techniques

The signal processing techniques apply a range of filters to an image, and the filter responses are analyzed for determination and classification of the texture of the image. The most common and direct approach for extracting texture features in this technique is applying any spatial domain filters such as Robert's masks, Laplacian masks or Law's filters to images (Canny, 1986; Laws, 1980a; Roberts, 1963). The spatial filters give strong responses to edge related features as the density of edges per unit area can characterize the fineness of textures. Law's mask also called as energy filters, was introduced by Laws (1979, 1980a) and it filters images with specific one-dimensional mask convolved with images to give strong responses to flats, edges, spots and ripples. Tuceryan (1994) used the spatial moments from spatial filters (Laws, 1980a) to generate moment-based features (moments from small windows of the image are used) for texture segmentation. In order to learn the effect of illumination and geometrical changes occurring in image texture, studies were performed where a set of filter banks were applied to the image in a spatial domain. The bank of filters has a set of spatial-frequency selective and orientation-selective linear filters to capture texture features at different scales. For example, Leung and Malik (2001) used 48 filters (two Gaussian derivatives at six orientation and three scales, eight LoG filters, and four Gaussian filters) and the responses of filters are gathered to form data vectors. These data vectors are then clustered using k-means algorithm to select images with same texture classes to form a row of output data called textons. Finally, each image is represented as a histogram of textons. Similarly, Varma and Zisserman

(2005) used a method called Maximum Response (MR) and used only 36 filters (two Gaussian derivatives at six orientations and three scales) and selected only the maximum responses from each scale (i.e. six responses) which reduced the memory requirement and offered better performance. However, Varma and Zisserman (2009) showed that filters could remove some relevant texture information and suggested that the local patches from an original image could give better performance.

The basic frequency domain filter involves applying Fourier transforms to the image for texture analysis. Fourier transforms are widely used for texture extraction due to its robustness to additive noises in images. Discrete Fourier Transforms (DFT) is usually used for image processing (Brigham and Yuen, 1978). DFT is a sampled Fourier Transform and therefore does not contain all frequencies forming an image, but instead a set of samples which is large enough to describe the spatial domain image completely. The number of frequencies corresponds to the number of pixels in the spatial domain image, i.e. the image in the spatial and Fourier domain are of the same size. As Fourier transforms are applied to whole images, it fails to extract local texture information leading to the necessity of spatial filters (Bovik et al., 1990). These filters will be computed based on local windowed size rather than a whole image, giving combined feature information of spatial/frequency properties of texture. The most common joint spatial/frequency filters are windowed Fourier filters (Short Term Fourier Transform, STFT) (Azencott et al., 1997) and multi-channel Gabor filters (Bovik et al., 1990; Haley and Manjunath, 1999).

Besides, wavelet transforms are one of the most common methods of texture analysis (Mallat, 1987), which can be utilized for characterizing texture properties at multiple scales (Rioul and Vetterli, 1991; Unser, 1995). The main advantage of Discrete Wavelet Transform (DWT) over Fourier transform is its ability to extract frequency and location information. Chang and Kuo (1993) developed a tree-structured wavelet transform instead of a typical pyramid structure to analyze the components located in the middle frequency channels. Similarly, a wide range of wavelet transforms was developed for various purposes like extracting rotation-invariant features using a Radon transform (Jafari-Khouzani and Soltanian-Zadeh, 2005), Gabor transform (Arivazhagan et al., 2006), linear regression model (Wang and Yong, 2008) for texture analysis.

### 2.6.4 Model Based Methods

Model-based texture analysis methods try to construct an image model which will be able to describe the texture and also synthesize textures as the image model parameters

capture essential texture qualities (Tuceryan and Jain, 1993). Primarily, most of the model-based approaches aim to interpret an image texture by using a generative, stochastic or fractal model to generate texture patterns (Materka et al., 1998). However, the computational complexity due to stochastic modelling parameters is a primary problem. Though fractal models are found to be useful in natural texture modelling and discrimination, they fail to find a suitable position in describing local image structures due to the inability of orientation selectivity (Chaudhuri and Sarkar, 1995; Materka et al., 1998; Pentland, 1984). The most well-known model approaches include the Random Field(RF) models like Markov Random Fields (MRF) (Cohen et al., 1991; Yousefi and Kehtarnavaz, 2011), Gibbs Random Field (GRF) models (Derin and Cole, 1986; Geman and Geman, 1984). MRF models can capture and extract spatial contextual information from an image as it assumes that the intensity of each pixel depends only on the intensities of neighbouring pixels.

The critical issue with model-based texture analysis approaches is the difficulty in determining how to chose the correct model for a specific texture image and how to map texture with the selected probability model (Zhang and Tan, 2002). Since these models require many parameters to be determined, which is not insignificant with large neighbourhood size, makes it less popular among texture analysis methods.

## 2.7 Uses of Texture Analysis

Texture analysis can be used for many applications in image processing and computer vision. The different purpose/problem domains are:- a) texture classification, b) texture segmentation, c) texture synthesis and d) texture shape analysis (Hau, 2015; Ojala et al., 1996; Zhou, 2006).

### 2.7.1 Texture Classification

Texture classification involves grouping the feature samples into different classes, where the samples in a class share some standard features, while each class differs on some features. No matter how good the feature extraction algorithm is in extracting features, the discrimination of texture regions depends also on an appropriate classification algorithm. Due to the complexity of feature space, the complete classification between classes is impossible, so classification depends on determining the probability for each class.

Texture classification can be performed in two stages: the learning phase and the recognition/discrimination phase. The texture contents extracted through texture de-

scriptors are learned through specific texture properties of image regions like similarity, fineness, homogeneity, and contrast. Once the model has been built with respective texture classes from the learning phase, it can be used to classify the (previously unseen) testing data, which should have the same properties as the training data. Subsequently, each unknown sample is compared with the training sample for the best similarity match using some classification algorithm (Ojala et al., 1996, 2000).

In general, for a classification process, selecting a portion of data from the feature set for testing and using the rest for training the classifier is known as *cross-validation*. This helps in ensuring the fidelity and non-randomness of reported classification outcome when a distinct sample is inspected. The major cross-validation approaches are *hold-out* method (feature data set is divided into two disjoint groups as training and testing set) and *bootstrap* method. The *K-fold* cross-validation method splits data into $K$-subsets where one of the folds is chosen for testing and rest for training. The extreme case of K-fold is considered to be the *leave-one-out* approach where the number of folds, $K$ is equal to the data set length, $N$. The bootstrap method, unlike the hold-out method, randomly generates sample sets of size, $N$ by sampling with replacement. So, if a data sample is chosen for training or testing, it cannot be drawn again.

### 2.7.2 Texture Segmentation

Image segmentation is the stage of texture analysis, where the group of pixels/region in image sharing similar texture characteristics like contrast, homogeneity, and fineness are estimated and grouped. The challenging aspect of segmentation is the effect by noise on texture features making it difficult to differentiate texture patterns. The main approaches to texture segmentation are:-

- Region-based approach: The method groups pixels in an image into homogeneous regions. The major region based categories are *Region-growing*, *Split and merge*, *Watershed transform*, etc.

- Contour-based methods: They focus on the detection of edges based on grey levels, colour, and texture for separating distinct regions. A few of them are Roberts edge detection, Sobel edge detection, Prewitt edge detection, Kirsch Edge detection, Robinson edge detection, and Canny edge detection.

- Clustering-based approaches: The pixels having same/similar properties are clustered into a group for segmentation. The most common approaches are

K-means clustering methods and Fuzzy C-means (FCM) methods, Mean shift (a non-parametric feature-space analysis technique for locating the maxima of a density function), Expectation maximization clustering, Hierarchical clustering (Jain et al., 1999) (creates a hierarchical tree of similarities between the vectors, called a dendrogram), Self-Organizing Map (SOM) (used to map a multi-dimensional dataset onto a (typically) two-dimensional surface), and Graph-theoretic approach.

- Thresholding methods: The most common, but faster and inexpensive way of segmentation, where a piece of global information about the image region, like histogram, is used to separate the regions. The main challenge with this technique is the proper decision of the threshold value and the ambiguity posed by different objects with a similar colour.

- Energy function-based methods: The most common methods in this approach are active contour/deformable models and Markov Random Field (MRF).

The image segmentation methods separate the whole image into subsections called regions thereby reducing the effects of additive noise caused by artefacts (Ali et al., 2019). Image segmentation plays a vital role in breast applications like localising suspicious regions, providing quantitative assessment and monitoring the progression of breast diseases as well as to analyse the anatomical structures (Saidin et al., 2012). Es-salhi et al. (2017) discussed a wide variety of mammogram image segmentation methods used in literature.

### 2.7.3 Texture Synthesis

Texture synthesis is the process of developing more substantial digital data from a small digital texture sample using the characteristics of the initial sample. It is a widely used technique in computer vision and image/video post-processing stages. The most common application is filling-up the holes in images as in inpainting or constructing large non-repetitive background images. The most common techniques used for texture synthesis are tiling (where multiple copies of the initial image are copied as tile, side by side), stochastic texture synthesis (it randomly chooses colour values of pixels based on parameters like minimum brightness, and average colour. It works better with stochastic texture images than with structure images), Chaos mosaic (developed by Microsoft group for internet graphics. It uses tiling method, and then random size patches of sample image are copied to output image followed by a filtering process

on the final image to smoothen the edges), Pixel-based texture synthesis (it copies pixels similar to the local neighbourhood to form a synthetic texture. The most typical methods used are Markov fields, non-parametric sampling, and tree-structured vector quantization), patch-based texture synthesis (generates a new texture image by copying and stitching together textures at various offsets using some clone tool. Image quilting and graph-cut are the most common approaches), Deep Learning and Neural Network approach (they are considered to be the fast and data-driven parametric approach).

Texture synthesis finds wide application in mammogram image processing. Developing synthetic images under controlled environment can help in developing better image quality and assessment. Bakic et al. (2002a) have developed a breast tissue pattern by simulating adipose tissue compartments. Later, Bakic et al. (2002a) developed a 3D model, based upon simulation of breast tissue structure and mammographic imaging process. Similarly, Bochud et al. (1999) developed a method called "clustered lumpy background" to generate synthetic medical texture images based on lumpy background and found that the first and second-order statistical features are similar with original and synthetic images. Different models were developed in literature to study and compare the inner tissue structure (Rose and Taylor, 2003), tumour pattern (Castella, 2009), abnormality, etc. in mammogram images.

### 2.7.4 Shape Analysis

Shape features are one among the fundamental characteristic texture properties. They helped in synthesizing the texture to develop 3D models and patterns. There are many cues in image texture that allow the viewer to infer the 3D shapes and surface of objects, like variations in shading, boundaries, and edges. The relationship between surface shape and variations in texture properties was studied by Gibson (1950). So, analyzing the surface geometry of an object from an image and inferring the 3D model of an object in space is vital in CAD applications. Similarly, Blostein and Ahuja (1989) used scaling effect to infer the surface information from images. The surface element extraction was performed in various scales by filtering the image with a filter at different scales, and the best surface fit was selected to define the size of the object. As a result, blob-like texture elements were extracted from different scales.

In medical imaging, especially in mammographic images, the shape estimation helps in better understanding of the nature of abnormalities, as malignant tumours may show distinctive shape and pattern when compared to benign cases. Additionally, it helps in simulating the 3D patterns of breast tissue and arriving at more inferences on the underlying tissue patterns and abnormalities.

## 2.8 Texture Feature Descriptors

Texture can be considered as an image property that explains the surface and structure of an image. So, texture feature descriptors can provide information like smoothness, coarseness, and regularity, which in turn can quantify the texture content of the object and optimize the appearance, shape, colour and size. This section will review a few texture descriptors and features commonly used in CAD systems, especially in this research work on mammogram images for texture analysis and classification.

### 2.8.1 First Order Feature Descriptors

Statistical features analyze the spatial distribution of grey level values at each point in the image by computing the local features at that point. Depending on the number of pixels defining the local texture feature, they can be first-order (one pixel), second-order (considering two pixels) or higher order. First-order statistics such as mean, standard deviation and higher-order moments of the histogram are concern with properties of pixels ignoring the spatial interaction between pixels. The most common method is the use of histogram for representing the grey-level distribution in image pixels.

Let random variable $I$ represents the gray levels of image region. The first-order histogram *P(I)* is defined as

$$P(I) = \frac{\text{number of pixels with gray level } (I)}{\text{total number of pixels in the region}} \tag{2.1}$$

Based on the definition of *P(I)*, the Mean $m_1$ and Central Moments $\mu_k$ of $I$ are given by

$$m_1 = E[I] = \sum_{I=0}^{N_g-1} IP(I) \tag{2.2}$$

$$\mu_k = E\left[(I - E[I])^k\right] = \sum_{I=0}^{N_s-1} (I - m_1)^k P(I), k = 2,3,4 \tag{2.3}$$

where $N_g$ is the number of possible gray levels (Hlaing, 2015; Kumar and Gupta, 2012). Commonly used central moments are Variance, Skewness and Kurtosis are given by $\mu_2$, $\mu_3$, and $\mu_4$, respectively. The Variance is a measure of the histogram width that measures the deviation of grey levels from Mean. Standard deviation is $\mu^{1/2}$ and is used to compute the dispersion of image data (pixel intensities). Skewness is the measure of the degree of histogram asymmetry around the Mean, and Kurtosis is a measure of the histogram sharpness (Hlaing, 2015).

### 2.8.2 Second-Order Feature Descriptors/GLCM Features

While first-order statistical features extract features based on individual pixels, the second-order statistical features compute the properties of two pixels according to their relative positions considering the neighbourhood relationship. Spatial grey level co-occurrence models are the most famous descriptor models that estimate the spatial image properties related to second-order statistics. Haralick (1979) developed one of the most popular second-order statistical texture feature descriptors using the grey-level co-occurrence matrices (GLCM). This method emphasizes the joint probability distribution of pairs of pixels. GLCM shows how often each grey level value occurs at a pixel located at a fixed geometric position relative to each other pixel, as a function of the grey level (Srinivasan and Shobha, 2008). GLCM represents how the pixel value, $i$ of a reference pixel occurs in a spatial relationship to a neighbouring pixel with pixel value $j$. So, each element $(i, j)$ of the matrix cell represents the number of occurrences of the pair of a pixel with pixel values $i$ and $j$ which are at a relative distance, d and orientation $(\theta = 0, 45, 90, 135)$ from each other as shown in Figure 2.2.

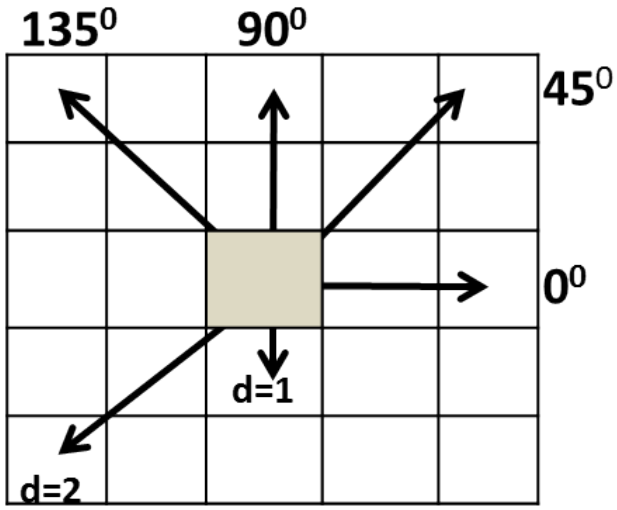



Fig. 2.2 An illustration of GLCM distance (d) and orientation ($\theta$), for extracting GLCM texture features

Computing GLCM features (known famously as Haralick features) involves constructing the GLCM and calculating features from the GLCM. Constructing GLCM involves creating a matrix based on the grey level values of the input image and constructing a grey level co-occurrence matrix using distance and orientation as shown in Figure 2.3.

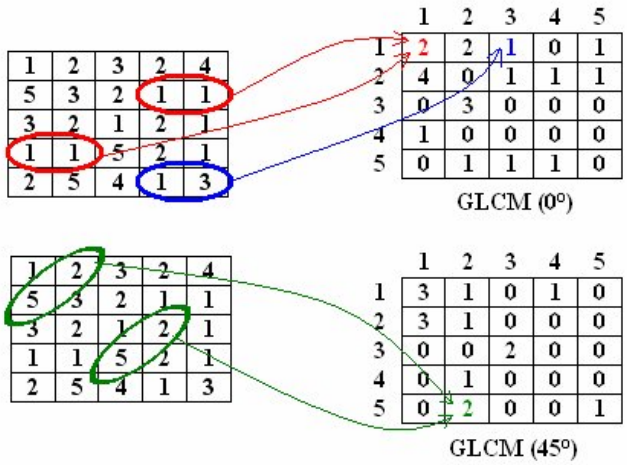



Fig. 2.3 Estimating GLCM matrix for distance (d=1) and orientation ($\theta = 0°, 45°$)

After the construction of GLCM matrices, the Haralick texture features (Haralick et al., 1973) are computed as follows:-

- *Notation:*

  where $p(i,j)$ : $(i,j)^{th}$ entry in a normalized grey-tone spatial-dependence matrix, $= P(i,j)/R$, ($R$ is a normalizing constant).

  $p_x(i)$ : $(i)^{th}$ entry in the marginal-probability matrix obtained by summing the rows of $p(i,j)$, $= \sum_{j=1}^{N_g} P(i,j)$.

  $N_g$ : Number of distinct gray levels in the quantized image.

  $\sum_i$ and $\sum_j$: $\sum_{i=1}^{N_i}$ and $\sum_{j=1}^{N_g}$, respectively.

  $p_y(j) = \sum_{i=1}^{N_g} p(i,j)$

  $p_{x+y}(k) = \sum_{\substack{i=1 \\ i+j=k}}^{N_g} \sum_{\substack{j=1 \\ i+j=k}}^{N_g} p(i,j), \quad k = 2,3,\cdots,2N_g$

  $p_{x-y}(k) = \sum_{\substack{i=1 \\ |i-j|=k}}^{Ng} \sum_{\substack{j=1 \\ |i-j|=k}}^{Ng} p(i,j), \quad k = 0,1,\cdots,N_g - 1$

- Energy: Energy, also called Angular Second Moment and Uniformity is a measure of textural Uniformity of an image. A homogeneous image contains a minimal number of grey tone transitions which result in fewer entries of larger magnitude resulting in considerable value for energy feature.

$$ASM/Energy = \sum_i \sum_j \{p(i,j)\}^2 \tag{2.4}$$

- Contrast: it is the estimation of the number of local variations present in an image.

$$Contrast = \sum_{n=0}^{N_{g-1}} n^2 \left\{ \sum_{\substack{i=1 \\ |i-j|=n}}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\} \tag{2.5}$$

where $n = |i-j|$ and $N_g$ is the number of grey levels in the image

- Correlation: It measures the linear dependency of grey levels of neighbouring pixels.

$$Correlation = \frac{\sum_i \sum_j (ij) P(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{2.6}$$

where $\mu_x,\ \mu_y,\ \sigma_x,\ \sigma_y$ are the means and std.deviations of $P_x$ and $P_y$, the partial probability density functions. where

$$\mu = \sum_i \sum_j i P(i,j) \tag{2.7}$$

and

$$\sigma = \sum_i \sum_j (i-\mu)^2 P(i,j) \tag{2.8}$$

- Sum of Squares/Variance: Variance is a measure of dispersion of data values around the mean with the combinations of reference and neighbour pixels.

$$Variance = \sum_i \sum_j (i-\mu)^2 P(i,j) \tag{2.9}$$

- Inverse Difference Moment: It measures the local homogeneity of image by measuring the closeness measure for the GLCM elements with GLCM diagonals. IDM achieves its largest value when most of the occurrences in GLCM are concentrated near the main diagonal. IDM is inversely proportional to GLCM contrast.

$$IDM = \sum_i \sum_j \frac{1}{1+(i-j)^2} P(i,j) \tag{2.10}$$

- Sum Average:

$$Sum\ Average = \sum_{i=2}^{2N_g} 2p_{x+y}(i) \tag{2.11}$$

where $x$ and $y$ are the coordinates (row and column) of an entry in the co-occurrence matrix, and $p_{x,y}(i)$ is the probability of co-occurrence matrix coordinates summing to x+y.

- Sum variance:

$$Sum\ Variance = \sum_{i=2}^{2N_g} (i - f_8)^2\, p_{x+y}(i) \tag{2.12}$$

- Sum Entropy :

$$Sum\ Entropy(f_8) = -\sum_{i=2}^{2N_g} P_{x+y}(i) \log\left\{P_{x+y}(i)\right\} \tag{2.13}$$

- Entropy :

$$Entropy = -\sum_i \sum_j p(i,j) \log(p(i,j)) \tag{2.14}$$

- Difference variance: It is variance of $P_{x-y}$

$$Difference\ Variance = \sum_{i=0}^{N_g-1} i^2 p_{x-y}(i) \tag{2.15}$$

- Difference Entropy:

$$Difference\ Entropy = -\sum_{i=0}^{N_g-1} p_{x-y}(i) \log\left\{p_{x-y}(i)\right\} \tag{2.16}$$

- Maximum Correlation Coefficient: It is the square root of the second largest Eigen value of **Q**, Where

$$\mathbf{Q}(i,j) = \sum_k \frac{p(i,k)p(j,k)}{P_x(i)P_y(k)} \tag{2.17}$$

- Info. Measure of Correlation$_1$:

$$Info.Measure\ of\ Correlation\ Feature_1 = \frac{HXY - HXY_1}{\max\{HX, HY\}} \tag{2.18}$$

- Info. Measure of Correlation $_2$:

$$Info.Measure\ of\ Correlation\ Feature_2 = (1 - \exp[-2(HXY_2 - HXY)])^{\frac{1}{2}} \tag{2.19}$$

where $HXY = -\sum_i \sum_j p(i,j)\log(p(i,j)), HX, HY$ are the entropies of $P_x$ and $P_y$ where

$$HXY_1 = -\sum_i \sum_j p(i,j)\log\left\{P_x(i)P_y(j)\right\} \tag{2.20}$$

and

$$HXY_2 = -\sum_i \sum_j P_x(i)P_y(j)\log\left\{P_x(i)P_y(j)\right\} \tag{2.21}$$

### 2.8.3 Tamura's Features

Tamura et al. (1978) developed a set of six features based on psychophysical studies of characterizing elements that are perceived in textures by humans. For an image $I$ with width $W$, height $H$, and the pixel at location $x$ and $y$ as $I(x,y)$, they defined the features as (Humeau-Heurtier, 2019):

- Contrast: It measures the variation of grey levels $q$: $q = 0, 1, ...q_{max}$ within the image region $g$ and to what extent their distribution is biased to black or white. The second-order (the variance $\sigma^2$) and normalised fourth-order central moments (kurtosis, $\alpha_4$) of the grey level histogram (empirical probability distribution) are:

$$T_{contrast} = \frac{\sigma}{(\alpha_4)^n} \tag{2.22}$$

where n=0.25 as suggested by Tamura for discriminating the textures.

$$\alpha_4 = \frac{\mu_4}{\sigma^4}; \quad \sigma^2 = \sum_{q=0}^{q_{\max}} (q-m)^2 \Pr(q|\mathbf{g}); \quad \mu_4 = \sum_{q=0}^{max} (q-m)^4 \Pr(q|\mathbf{g}) \tag{2.23}$$

and $m$ is the mean grey level, i.e. the first order moment of the grey level probability distribution.

- Directionality: It takes the details of edge strength and the directional angle. They are calculated using pixel-wise derivatives according to Prewitt's edge detector.

- Coarseness: It measures the scale of a texture. For a window size, a texture with a smaller number of texture elements is considered more coarse than with a larger number.

$$\text{Coarseness}(I) = \frac{1}{W \times H} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} S_{\text{best}}(x,y) \tag{2.24}$$

where $S_{\text{best}}(x,y) = 2^k$. $k$ is the value that maximizes $\max_{1 \le k \le L} \left(E_{k,h}(x,y), E_{k,v}(x,y)\right)$ in the horizontal or vertical direction, $k \in [1,L]$ where $2^L \le \min(W,H)$.$E_{k,h}(x,y)$ and $E_{k,v}(x,y)$ are computed from $A_k(x,y)$ which is the average at each pixel $I(x,y)$ of the neighborhood of size $2^k \times 2^k$ :

$$A_k(x,y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} \frac{I(i,j)}{2^{2k}} \tag{2.25}$$

From $A_k(x,y), E_{k,h}(x,y)$ and $E_{k,v}(x,y)$ are computed as

$$E_{k,h}(x,y) = \left| A_k\left(x+2^{k-1},y\right) - A_k\left(x-2^{k-1},y\right) \right| \tag{2.26}$$

and

$$E_{k,v}(x,y) = \left| A_k\left(x,y+2^{k-1}\right) - A_k\left(x,y-2^{k-1},y\right) \right| \tag{2.27}$$

- Linelikeness: It is defined as the average coincidence of edge directions that co-occur at pixels separated by distance '$d$' along the direction $\alpha$.

- Regularity: It is defined as the

$$Regularity = 1 - r\left(\sigma_{\text{coarseness}} + \sigma_{\text{contrast}} + \sigma_{\text{directionality}} + \sigma_{\text{linelikeness}}\right) \tag{2.28}$$

where $r$ is the normalizing factor and $\sigma$, the standard deviation of the feature in each sub-image of the texture.

- Roughness: It is the sum of the coarseness and contrast measures.Roughness is related to the standard deviation of the normalized grey levels.

### 2.8.4 Local Binary Pattern Methods

Local Binary Patterns (LBP) is a non-parametric texture descriptor whose aim is to efficiently summarize the local structures of images developed by Ojala et al. (1996). It is computationally simple and is invariant to monotonic grey scale changes and is convenient for multi-scale feature extraction. LBP operator describes the neighbourhood pixels of a central pixel by generating a bit-code depending on the central pixel value. Later, a decimal value is computed out of the bit code for each central pixel. The histogram is calculated to extract the features from the LBP pattern image. The detailed description of LBP and its variants for breast density estimation is explained in Chapter 3.

### 2.8.5 Local Directional Pattern

Jabid et al. (2010) developed a local directional feature descriptor for facial feature recognition. The Local Directional Pattern (LDP) feature is obtained by computing the edge response value of each pixel in eight different directions. Thereafter a code is generated from the relative strength magnitude for each pixel. The detailed explanation of the LDP feature detector will be covered in Chapter 3.

## 2.9 Datasets

Access to medical data is not an easy task due to ethical issues relating to information about the patient, security and privacy. The use of a standard database assists in standardizing the evaluation comparison of various CAD systems and algorithms. The datasets used in the thesis are the Mammographic Image Analysis Society (MIAS) and the Digital Database for Screening Mammography (DDSM).

### 2.9.1 MIAS

The Mammographic Image Analysis Society (MIAS) is an organisation of UK research groups interested in the understanding of mammograms and has generated a database of digital mammograms. The X-ray films were taken from the UK National Breast Screening Programme. The images have been digitised to 50-micron pixel edge with a Joyce-Loebl scanning microdensitometer, representing each pixel with an 8-bit word. The database contains 322 digitised films including 208 normal, 63 benign and 51 malignant mammograms. It also includes radiologist's ground truth markings on the locations and type of any abnormalities that may be present along with the details

on tissue density type (fatty, fatty-glandular or dense-glandular). The database is also available as mini-MIAS, which has been reduced to a 200-micron pixel edge, and padded/clipped making all images of size 1024x1024 (it should be noted that mini-MIAS has not been used in this thesis). Details on the MIAS database can be found in (Davies, 1993). The dataset includes details about the character about background tissue (fatty, fatty-glandular, dense glandular), class of abnormality present (CALC-Calcification, CIRC-well-defined/circumscribed masses, SPIC-Spiculated masses, MISC-ill-defined masses, ARCH-architectural distortion, ASYM-asymmetry, NORM-normal), the severity of abnormality (benign/malignant), image-coordinates of abnormality, the approximate radius of a circle enclosing the abnormality.

In the thesis, the tissue density type, BI-RADS density class information, and the microcalcification class type from MIAS have been utilised for evaluating the performance of the proposed algorithms. Each mammogram is annotated into three distinct groups as either Fatty (F), Fatty-Glandular(G) or Dense-Glandular (D) based on the tissue pattern. We perform classification on 321 images (it should be noted that a single image could not be included as it had been corrupted and the downloaded database only contained 321 images) with 106 fatty cases, 104 fatty-glandular and 111 dense-glandular cases. We classify the mammograms into three categories- Fatty (F), Glandular (G) and Dense (D) based on the tissue structure. More details on how the images are used will be explained in respective chapters/sections.

### 2.9.2 DDSM

Digital Database for Screening Mammography (DDSM) is a publically available mammographic database used by many researcher communities (Heath et al., 2000). The database and its details are provided at http://www.eng.usf.edu/cvprg/Mammography/Database.html. It was developed as a part of collaborative work between Massachusetts General Hospital, Sandia National Laboratories and the University of South Florida Computer Science and Engineering Department. The dataset contains approximately 2,500 studies where each study includes two images of each breast along with a bit of patient information (like age, and ACR density rating) and image information (like scanner used, and spatial resolution). A pixel-level 'ground truth' information about the location and type of suspicious regions in the images are provided with the images. For each woman, four mammograms were taken, including both MLO and CC views of each breast. The scanners used for mammogram digitisation included DBA M2100 ImageClear (42 microns per pixel, 16 bits), Howtek 960 (43.5 microns per pixel, 12 bits), Lumisys 200 Laser (50 microns per pixel,

12 bits), and Howtek Multi- Rad850 (43.5 microns per pixel, 12 bits). Each mammogram included patient information, breast density rating (four BIRADS density categories), and subtlety rating for abnormalities. A detailed description of the scanner, cases, volumes, software to access the mammogram and truth images are provided in 'http://www.eng.usf.edu/cvprg/Mammography/Database.html' (Rose et al., 2006).

In the thesis, we have used the BIRADS density information for a set of mammograms are used for breast density estimation and microcalcification classification. Detailed information on images will be described in respective chapters.

### 2.9.3 OPTIMAM

The Optimam Mammography Image Database (OMI-DB) has been developed to support researchers involving medical imaging. The database is under an ongoing project at the Medical Physics Department of the Royal Surrey County Hospital, which includes NHS Breast Screening Programme (NHSBSP) images from different screening centres across the United Kingdom with an aim to develop a large repository of breast images for research purposes. The database contains 2D and 3D processed and unprocessed set of breast images, associated annotations which describe features of abnormalities such as microcalcification, mass, architectural distortions. The images were categorized by radiologists into three clinical categories: normal, benign, and malignant cases. These mammograms were acquired using a Hologic Selenia mammography unit, with a resolution of 70 microns per pixel and a depth of 12 bits (Halling-Brown et al., 2014). More information on the database can be seen from 'http://commercial.cancerresearchuk.org/optimam-mammography-image-database-and-viewing-software'.

## 2.10 Summary

The chapter provided an idea about the need for CAD systems and their support for radiologists. Besides, a detailed study of breast density and microcalcification classification was done. Moreover, a brief introduction to texture, its definition and application in computer vision and image processing were provided. Besides, various uses of texture analysis techniques along with the most commonly used feature descriptors are explained in detail as they will be used in the following chapters for texture feature analysis. Finally, the databases used for the research work are explained in detail.

# Chapter 3

# Breast Density Classification using LBP variants

Following the previous chapter on texture analysis, in this chapter, we discuss the nature of local features for texture extraction. The focus of the discussion is to address the following aspects: utilizing Local Binary Patterns (LBP) and its variants for feature extraction and density classification of mammogram images; what are the effects of adding intensity attributes to texture features for mammogram density classification. Towards the end of chapter, a new LBP variant, named Mean-Elliptical Local Binary Pattern (M-ELBP), which incorporates the intensity features along with the texture is introduced. In addition, the effect of Region of Interest (ROI) size and location, the effect of descriptor size, and comparison between ROI and whole breast for density classification are explored in this chapter.

## 3.1 Background Work

Texture can be considered as the repetition of sub-patterns/local patterns or microstructure with a number of local properties generating a global structure/pattern as shown in Figure 3.1. However, this property makes texture not defined at a single point in an image as it is considered to be repeating patterns of local variations in an image.

A common problem with texture patterns is that they are not uniform due to the illumination, scale, and variation in orientation. Ojala et al. (1996) found LBP to be robust to monotonic gray-scale changes and is considered as a potentially powerful gray-scale invariant texture descriptor.

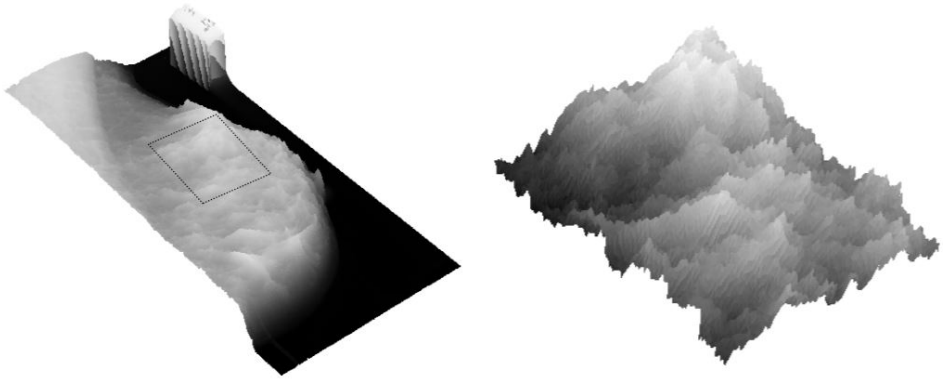


Fig. 3.1 Representation of texture for an imahe region, (the image was taken from http://www.cyto.purdue.edu/cdroms/micro2/content /education/wirth06.pdf).

## 3.2 Local Binary Patterns and its Applications

Due to its simplicity and efficiency in local feature extraction, the LBP was considered as one of the best methods for local feature extraction. It can be considered as a unifying approach of statistical and structural feature extraction with low computational complexity. The possibilities of expanding and modifying the algorithm, like the neighborhood topology size and multiscale feature extraction paved a way for developing different variants of LBP for numerous applications. In particular for texture classification (Cheung and Deng, 2014), face recognition (Lu et al., 2018; Xiaohua et al., 2017), fingerprint processing (Al-Nima et al., 2017; Zaghetto et al., 2017) to medical image analysis (Abbasi and Tajeripour, 2017; Babaie et al., 2017).

### LBP in Medical Image Processing

Nanni et al. (2010) conducted experiments on biomedical images for different applications based on morphological variants of LBP. The experiments showed that the elongated quinary patterns (EQP) based on an elliptic neighbourhood with five scale for encoding the local grey-scale difference performed better for extracting texture features. Nanni et al. (2010) also used an Elongated Ternary Pattern (ELTP) feature extractor and produced a reliable system for classifying pain states. Oppedal et al. (2015), used LBP features for classifying dementia from different regions of magnetic resonance imaging (MRI) images. Unay et al. (2007), showed that LBP is invariant to some of the MRI artefacts. Mukundan (2018) used LBP for Human Epidermal Growth factor Receptor 2 (HER2) protein scoring for breast cancer histology slides. Similarly, image retrieval methods played an important role in medical applications.

LBP and its variants have found a prominent role in texture feature extraction in image retrieval. Xu and Zhang (2009) developed LBP with Image Euclidean Distance (IMED) for feature extraction and compared it with other variants like Improved Local Binary Pattern (ILBP), Extended Local Binary Pattern (ELBP), Local Gabor Binary Pattern (LGBP) for image retrieval with an accuracy of 92.5% for IMED. Ko et al. (2011) used random forests with proposed local wavelet-based local binary pattern (LBP) to classify x-ray images. Wan et al. (2017) used integrated LBP features for classifying breast tissue by optical coherence microscopy (OCM) images.

Some of the other applications using LBP and its variants are described in Table 3.1 to show the potential of LBP texture descriptor in feature extraction.

## 3.3 Local Binary Pattern

LBP, first introduced by Ojala et al. (1996), had low computational complexity and low sensitivity to illumination changes. Ojala et al. (1996) introduced an LBP operator based on the assumption that the texture of an image region has two aspects, the pattern, and its strength. In its simplest form, LBP considers the $3 \times 3$ surrounding of a pixel (forming a circle of radius $R$=1). The central pixel is compared with each neighboring pixel ($i = 8$) to generate a binary pattern. The pixels are followed along a circle, i.e. clockwise or counterclockwise direction. If the central pixel, $g_c$ is greater than the neighbouring pixel ($P_i$) under consideration ($x = P_i - g_c$), a binary '0' is given to neighbouring pixel else a binary '1' is assigned. This gives an 8-digit binary number to the central pixel. Later the binary pattern is converted to a decimal number for convenience. The histogram is computed over the cell to estimate the frequency of each number occurring. The histogram can be seen as a 256-dimensional feature vector for an 8-bit pattern ($2^8 = 256$). The histograms are concatenated for all cells to give the feature vector for the whole image. A pictorial representation of LBP estimation is shown in Figure 3.2.

A detailed calculation of an LBP pattern is provided through equations 3.1 - 3.4.

$$C = \begin{pmatrix} 70 & 80 & 105 \\ 20 & 100 & 110 \\ 20 & 10 & 100 \end{pmatrix} \tag{3.1}$$

Table 3.1 An overview of LBP Variants and its applications

| LBP Variants | Description | Reference | Applications |
|---|---|---|---|
| Local Gabor Binary Pattern Histogram Sequence (LGBPHS) | Gabor filtering applied to the image before LBP computation | (Zhang et al., 2005) | Face Representation and Recognition |
| Elongated LBP (ELBP) | Elongated neighbourhood distribution was used | (Liao and Chung, 2007) | Face Representation and Recognition |
| Elongated Quinary patterns (EQP) | Elongated neighbourhood distribution was used with quinary encoding (-2,-1,0,1,2) for patterns | (Liao and Chung, 2007) | Medical image analysis |
| Improved LBP (ILBP) | Considers both local shape and texture information, estimate the neighbourhood mean value as the threshold | (Jin et al., 2004) | Face detection |
| Fuzzy Local Binary Pattern (FLBP) | Fuzzification allows contributing to more than a single bin in the distribution of the LBP values used as a feature vector | (Iakovidis et al., 2008) | Ultrasound Texture Characterization |
| Probabilistic Local Binary Pattern (PLBP) | PLBP is able to encode the magnitude of the difference between two compared pixels by employing probability, combined with Linear Discriminant Analysis (LDA) to extract its most discriminant information | (Tan et al., 2009) | Ultrasound Texture Characterization |
| Elongated Ternary Pattern (ELTP) | Three values are used (-1,0,1) | (Nanni et al., 2010) | Classifying pain states, Medical image processing |
| Bayesian Local Binary Pattern (BLTP) | Labelling is modelled as a probability and optimization process | (He et al., 2008a) | Texture retrieval |
| Pyramid-Based Multi-structure Local Binary Pattern | Executing the LBP method on different layers of an image pyramid | (He et al., 2010b) | Texture retrieval |
| LBP variance (LBPV) | Build rotation variant LBP histogram and then apply a global matching | (Guo et al., 2010a) | Texture analysis and global matching |
| Adaptive LBP (ALBP) | Stable directional statistical features | (Guo et al., 2010b) | Texture analysis |
| Local Derivative Pattern (LDP) | nth-order LDP is proposed to encode the (n-1)th -order local derivative direction variations | (Zhang et al., 2010) | Face identification and face verification |
| Local Ternary Pattern (LTP) | Three values (-1,0,1) used for creating bit pattern | (Tan and Triggs, 2010) | Face recognition |
| Enhanced Local Ternary Pattern (ELTP) | Three values (-1,0,1) used for creating bit pattern,auto-adaptive strategy for threshold selection is adopted | (Yuan et al., 2014) | Texture classification |

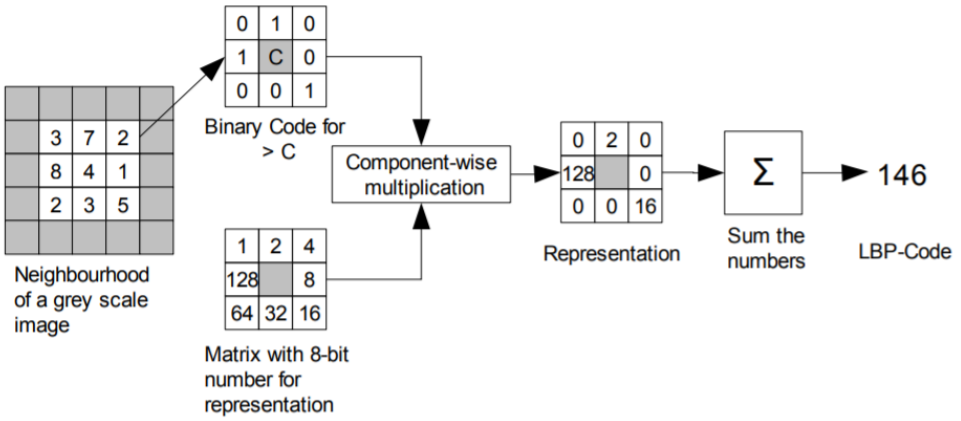



Fig. 3.2 Representation of LBP calculation for a pixel (the image taken from Lindahl (2007)).

C is a $3 \times 3$ grey scale block from an image with the central pixel at location (i, j), then the LBP code after applying the comparison is determined by $s$, where

$$s(x) = \begin{cases} 1, & \text{for } x \geq 0. \\ 0, & \text{for } x < 0. \end{cases} \tag{3.2}$$

Finally, the 8 bit LBP pattern is calculated using

$$LBP = \sum_{p=0}^{7} s\,(g_p - g_c)\,2^p \tag{3.3}$$

$$\begin{pmatrix} s(70-100) & s(80-100) & s(105-100) \\ s(20-100) & 100 & s(110-100) \\ s(20-100) & s(10-100) & s(100-100) \end{pmatrix} \xrightarrow[\text{Pattern}]{\text{Binary}} \begin{pmatrix} 0\times2^7 & 0\times2^0 & 1\times2^1 \\ 0\times2^6 & & 1\times2^2 \\ 0\times2^5 & 0\times2^4 & 1\times2^3 \end{pmatrix} \xrightarrow[\text{Decimal}]{LBP_{(8,1)}} 14 \tag{3.4}$$

A sample of LBP estimation for a pixel is shown in Figure 3.2 followed by histogram computation for each cell in the image and histogram concatenation for the whole image is shown in Figure 3.3.

### 3.3.1 LBP Variants

A wide variety of LBP variants were developed by researchers by updating certain parameters like the neighbourhood topology, pre-processing methods before LBP calculation, thresholding and encoding variations, multiscale analysis and feature

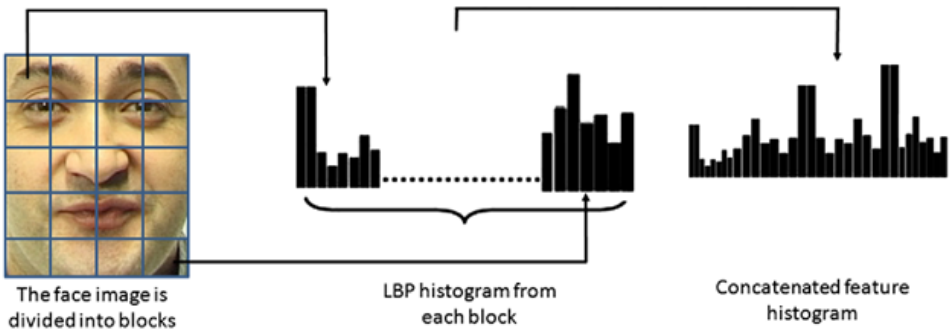



Fig. 3.3 LBP histogram estimation.

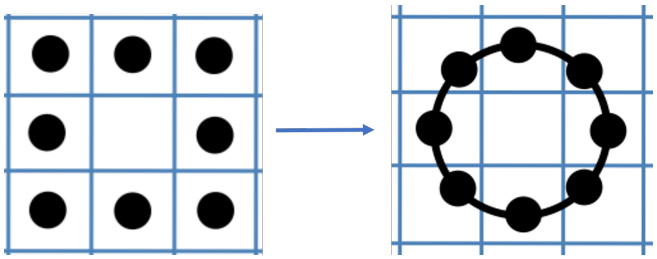


Fig. 3.4 LBP eight pixel neighbourhood. Square window (left) vs circular (right) neighbourhood.

estimation,and handling rotations. The following Section discuss LBP variants related to the current work:-

### *Circular LBP*

This is the classical LBP in which the neighborhood for each central pixel is selected as eight neighbors in a local window. It is clear that the four neighbors (at the corners) are at a distance of $\sqrt{2}$ compared to the other pixels which have a unity distance to the central pixel. To avoid this problem, a circular neighborhood was introduced as shown in Figure 3.4 maintaining each neighbor pixel at equal distance from the central pixel (Ojala et al., 2002). In practice, the neighboring pixels are sampled on a circle, such that the grey values of the neighborhood which are not exactly in the center of pixels are estimated by interpolation. So LBPs can be defined as a circular symmetric neighborhood with $R$ radius distance and $P$ neighboring pixels. For a central pixel $g_c$, the neighboring pixels $g_p$ can be calculated using equation 3.5

$$(-R\sin(2\pi p/P) \; and \; (-R\cos(2\pi p/P) \tag{3.5}$$

where $p = 1,2,...P$

Despite the advantages, the classic LBP have significant disadvantages (Liu et al., 2017) like

- Generating long histograms, even for small neighborhoods leading to larger storage requirements.
- Inability to detect large-scale textural structures though it was excellent in capturing the local structures.
- Highly sensitive to image rotation.
- Highly sensitive to noise. Due to the strict classification of the threshold, even the noise value are considered as contrast.
- High chance of losing local textural information due to the use of hard, fixed, and coarse quantization methods.

To overcome the problems of basic LBP descriptor, Ojala et al. (2002) observed that certain LBP patterns represent the fundamental texture microstructures, and named these as uniform patterns. The rotation invariant and uniform LBPs not only reduced the feature length of traditional LBP but also worked better on rotated texture data.

1. **Uniform LBP:**

   Certain binary patterns are prominent and appear more frequently (called uniform patterns) in an image compared to other patterns. These uniform patterns were retained leading to uniform LBP ($LBP^{u2}_{r,p}$). In general, the uniformity measures and counts the number of 0/1 or 1/0 transitions between successive bits in the circular representation of the pattern binary code as:

   $$U\left(\mathrm{LBP}_{r,p}\right) = \sum_{i=1}^{p} \left| s\left(g_{mod(i,p)} - g_c\right) - s\left(g_{i-1} - g_c\right) \right| \tag{3.6}$$

   If the number of transitions U is greater than 2, it is called a nonuniform pattern and is considered as a single group. An example of a non-uniform and uniform pattern is shown in Figure 3.5 and the list of possible uniform LBP patterns are shown in Figure 3.6, where black dot represents '0' and white represents '1'.

2. **Rotation Invariant Uniform LBP:**

   In order to improve rotation invariance and reduce feature dimensionality, Ojala et al. (2002) proposed the rotation invariant uniform LBP descriptor ($LBP^{\mathrm{riu2}}_{p,r}$)

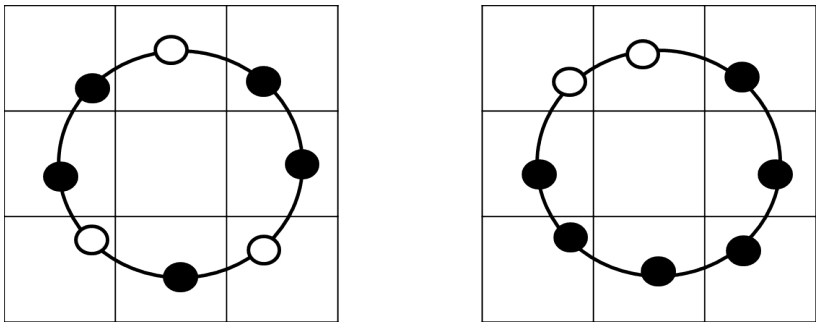


Fig. 3.5 LBP non-uniform pattern (left) and uniform pattern (right).

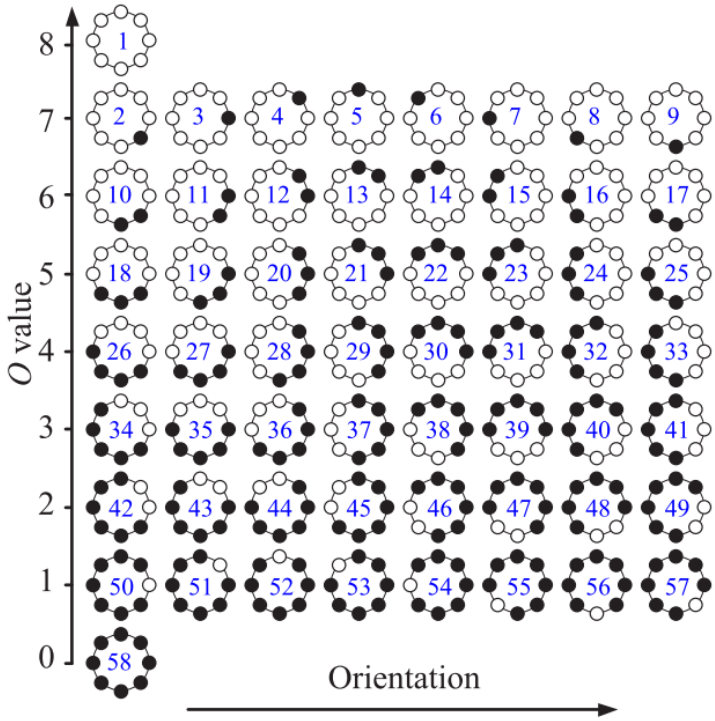



Fig. 3.6 List of Uniform LBP Patterns (image courtesy Liu et al. (2017)).

$$LBP_{p,r}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP_{p,r}) \leq 2 \\ p+1 & \text{otherwise} \end{cases} \tag{3.7}$$

and this can be calculated using equation 3.7. To compare, while LBP generated 256 binary patterns for a neighbourhood of 8, $LBP^{u2}$ generated 58 uniform patterns while $LBP^{riu2}$ generated only 9 patterns. An example of which is shown in Figure 3.7


Fig. 3.7 LBP patterns for rotation invariant uniform LBP for a neighbourhood of eight.

### Multi-scale LBP

It is the method of texture feature extraction using an LBP descriptor by defining different radius around a central pixel. In this way, features around a pixel at multiple scales can be calculated. Thereafter, the histograms obtained at different scales are concatenated to create the feature vector. This can result in a high dimensional feature vector. A sample of multi-scale LBP is shown in Figure 3.8.

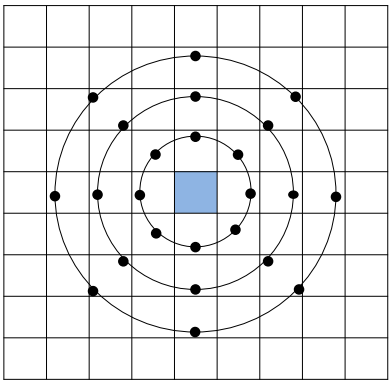


Fig. 3.8 LBP patterns for multiscale LBP with eight-neighbourhood with different radius.

### Local Ternary Pattern (LTP)

Unlike the LBP descriptor, the LTP descriptor (Tan and Triggs, 2010) does not threshold the pixels into 0 and 1. In LTP, a 3-value (-1,0,1) code is suggested called a local

ternary pattern, with an encoding interval of $\pm t$ around the centre pixel as

$$s(g_p, g_c, t) = \begin{cases} 1 \text{ if } g_p \geq g_c + t \\ 0 \text{ if } \left|g_p - g_c\right| < t \\ -1 \text{ if } g_p \leq g_c - t \end{cases} \tag{3.8}$$

where, $g_p$ is the neighbouring pixel of central pixel $g_c$, and $t$ is a user-specified threshold. Compared to LBP, which tends to be sensitive to noise due to a hard classification of 0 or 1, LTP codes are more resistant to noise. But, it is not strictly invariant to grey level transformation. Each ternary pattern is then separated into two different LBP patterns (both lower patterns and upper patterns). Any values of -1 for the upper pattern get mapped to 0 and 1 for the lower pattern. Similarly, for estimating the lower pattern, any values that were 1 in the original window get mapped to 0. An example of pattern generation for the LTP descriptor is shown in Figure 3.9. Thereafter, separate histograms are computed for each pattern and then concatenated which leads to double the size of LBP feature vector. LTP and its variants like completed LTP (CLTP) have found many applications for texture classification (Yuan et al., 2014), and face recognition (Ibrahim et al., 2014).

### 3.3.2 Topology based LBP Variants

Traditional LBP variants focus on circular topology by varying the neighborhood size, radius size or exploring multi-scale features. But a great variety of neighborhood topology could be exploited to overcome the limitations of traditional LBP. Though the circular neighborhood is good for rotation invariance, there are some applications like face recognition where anisotropic information has more importance than rotation invariance.

The basic circular variants like Elliptical Binary Patterns (ELP) (Liao and Chung, 2007; Nanni et al., 2010), have found wide applications in face recognition and medical image analysis. More variants like parabolic, hyperbolic, and spiral neighborhood topologies were introduced by Nanni et al. (2010) for identifying pain states from face features as shown in Figure 3.10.

Similarly, Petpon and Srisuk (2009) developed a local line binary pattern where the binary codes are obtained in vertical and horizontal directions separately along with

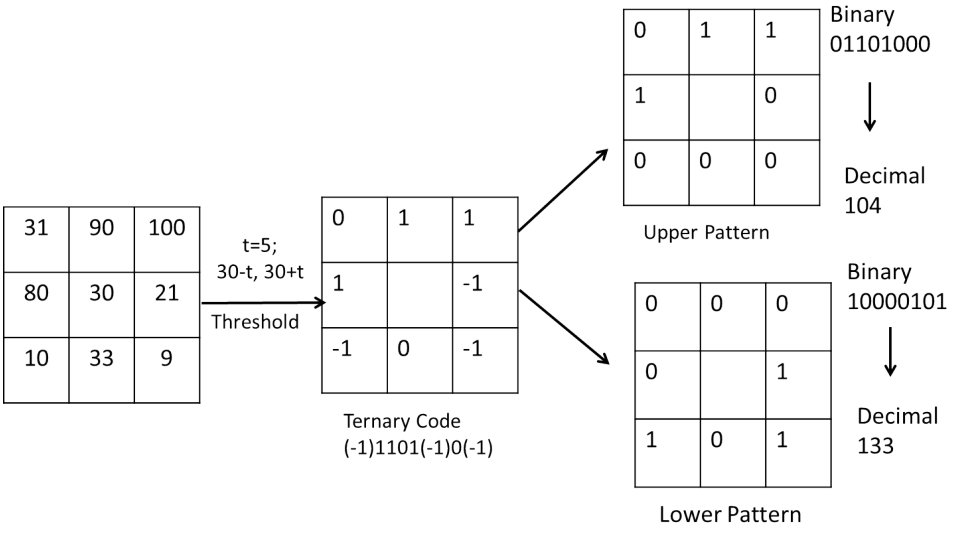



Fig. 3.9 LTP pattern generation with a threshold $t = 5$.

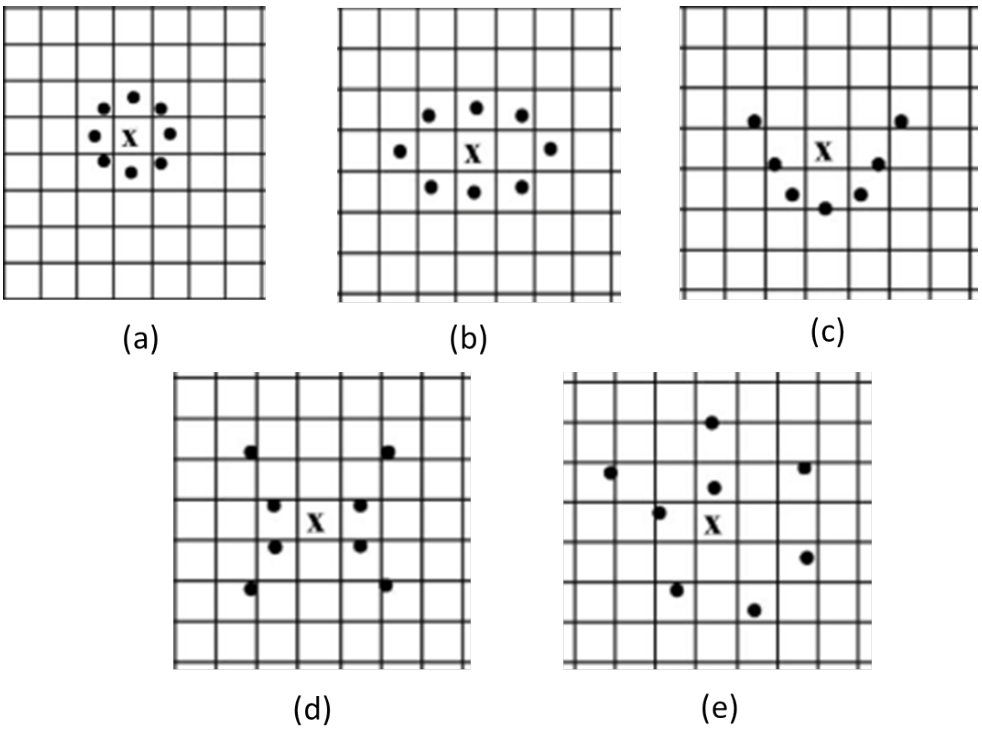



Fig. 3.10 Elliptical topology variants (image courtesy (Nanni et al., 2010)).

its magnitude characterizing the changes in intensity like edges and corners. Likewise, Local Quantized Pattern (LQP) (Hussain et al., 2012) utilized the lookup tables and vector quantization techniques to analyze the local features in detail, using larger sized neighborhood. The circular neighborhood is an advantage for a texture descriptor but there are applications where the anisotropic texture information has more potential in distinguishing the objects. The circular-like topologies like ellipse help in attaining more features from different directions by rotating at different orientations (Liu et al., 2017). In order to extract more micro-pattern features from the extracted ROIs of a mammogram, we introduce a new texture feature extractor, Mean-Elliptical Local Binary Pattern (M-ELBP) which is more robust to noise compared to traditional LBP and could extract more directional features due to its elliptical structure.

## 3.4 Mean Elliptical Local Binary Pattern

As explained in previous sections, Ojala et al. (2000) proposed LBP for each central pixel $(x_c, y_c)$ in the input image with a grey level value $g_c$. It's LBP value is estimated following the pixels along a circular path either clockwise or counter-clockwise as in Figure 3.11. The histogram for the whole ROI was calculated to extract the histogram texture features. The $LBP^{p,r}$ was calculated as follows:

$$LBP^{p,r}(x_c, y_c) = \sum_{i=1}^{p} s(g_i^{p,r} - g_c)2^{i-1} \tag{3.9}$$

where s(x) is defined as

$$s(x) = \begin{cases} 1, & \text{if } x >= 0 \\ 0, & \text{if } x < 0 \end{cases} \tag{3.10}$$

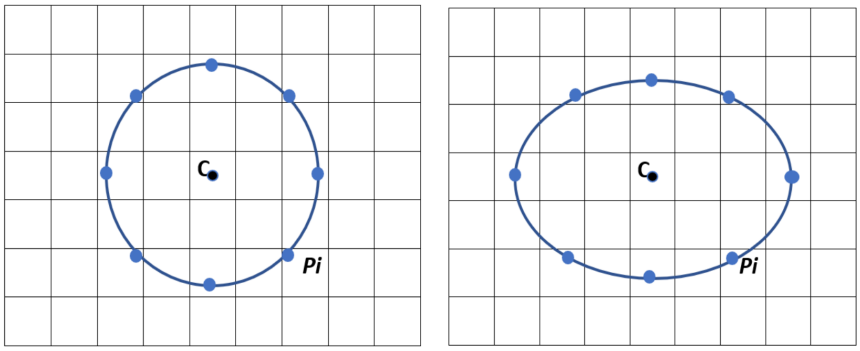
Using ELBP, for each central pixel $(x_c, y_c)$, we consider the neighbouring pixels $P$ which lies on an ellipse (with radius $R_1$ and $R_2$) which was calculated as :

$$ELBP^{P,R_1,R_2}(x_c, y_c) = \sum_{i=1}^{P} s(g_i^{P,R_1,R_2} - g_c)2^{i-1} \tag{3.11}$$

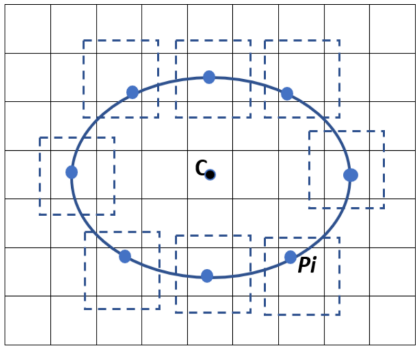
where the $i^{th}$ neighbouring pixel coordinate of $(x_c, y_c)$, is computed as :

$$angle_{step} = 2(\frac{\pi}{P}) \tag{3.12}$$

$$x_i = x_c + R_1(\cos((i-1)(angle_{step}))) \tag{3.13}$$




(a) LBP Pattern (b) ELBP Pattern




(c) M-ELBP Pattern

Fig. 3.11 LBP, ELBP, M-ELBP Patterns with neighbourhood of eight.

$$y_i = y_c - R_2(\sin((i-1)(angle_{step}))) \tag{3.14}$$

For M-ELBP computation, we consider the mean intensity values ($\bar{g}_i$) around each neighbouring pixel $P_i$ that lies on an ellipse (with radius $R_1$ and $R_2$) as

$$M\text{-}ELBP^{P,R_1,R_2}(x_c,y_c) = \sum_{i=1}^{P} s(\bar{g}_i^{P,R_1,R_2} - g_c)2^{i-1} \tag{3.15}$$

The mean intensity value around each neighborhood pixel (a local window of size $3 \times 3$) is compared with the central pixel to create a binary pattern as in LBP or ELBP.

Compared to the LBP patterns, the ELBP descriptors help in extracting more specific spatial features from the mammographic images as it could extract more directional features at different orientations encountering additional micro patterns and reducing the noise effect.

### 3.4.1 Feature Extraction using M-ELBP

In order to extract additional features from the mammographic images, we extend ELBP/M-ELBP descriptors at eight different orientations generating an eight-bit pattern for each central pixel. When $(R_1 = R_2)$, the ELBP/M-ELBP reduces to LBP descriptor while $(R_1 < R_2)$, we have a vertical ellipse and if $(R_1 > R_2)$ we get a horizontal ellipse structure. This will help in capturing more intrinsic features, in detail, from the mammograms. To extract additional intrinsic features, we modified the ELBP operator to the mean-ELBP (M-ELBP) where the intensity features at different orientations were attained along with the texture features. The detailed overview of the method is described in Figure 3.11. Compared to the LBP descriptor, the elliptical descriptors could extract more features when applied to ROIs at different orientations $\theta = \{0°, 45°, 90°, 135°, 180°, 225°, 270° \text{ and } 315°\}$. Figure 3.12 shows the effect of horizontal and vertical M-ELBP on the extracted image ROI for $P$=8 at an orientation $\theta = 0°$. The procedure is repeated for all orientations to extract M-ELBP pattern features. Figure 3.13 summarizes the classification of breast tissue using the ELBP/M-ELBP descriptor variants at eight different orientations with an elliptical radius of $(R_1 = 4, R_2 = 7)$ for the ELBP/M-ELBP descriptor for a neighborhood range of eight pixels. Subsequently, the histogram labels for the ELBP/M-ELBP descriptors at all orientations are concatenated to generate the texture feature vector for the ROIs.

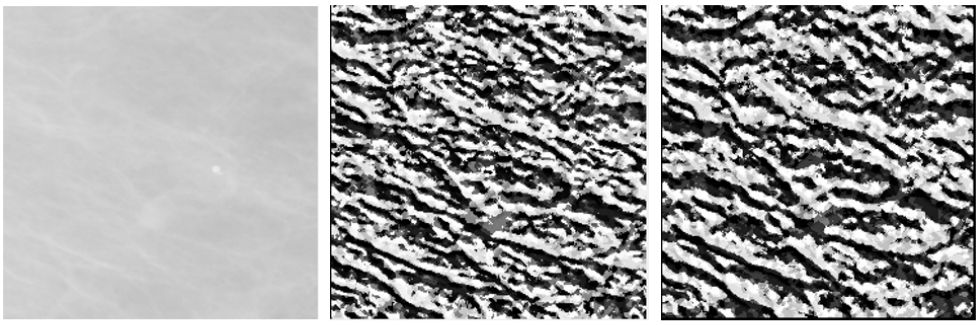



Fig. 3.12 Effect of horizontal and vertical M-ELBP on mammographic RoI (a) Original Image (b) Vertical M-ELBP pattern for $M-ELBP^{8,7,4}$ (c) Horizontal M-ELBP pattern for $M-ELBP^{8,4,7}$.

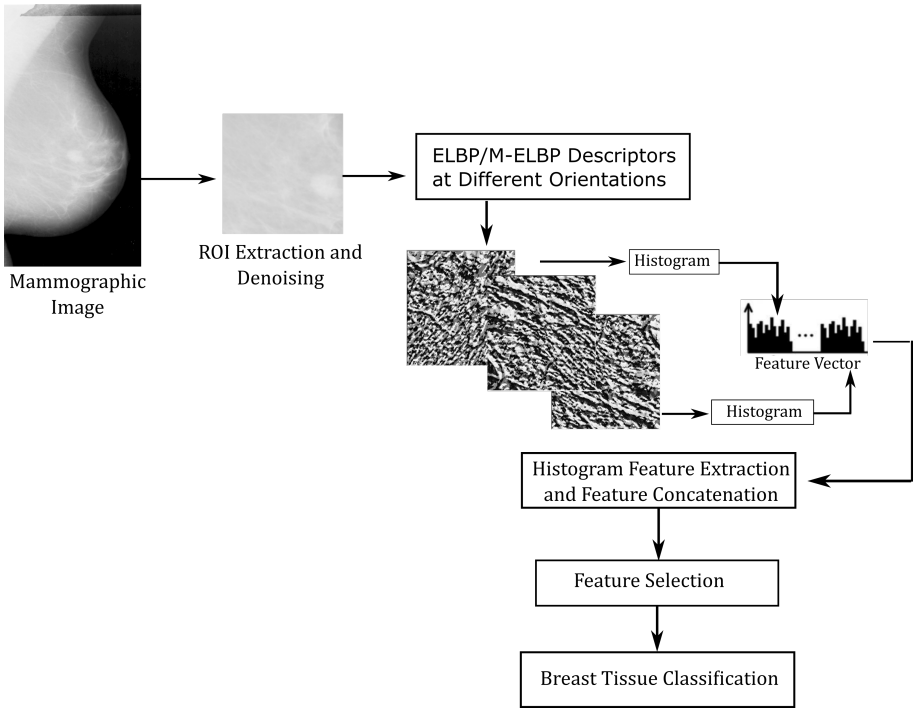



Fig. 3.13 Summary of ROI selection, feature extraction and classification using ELBP variants.

## 3.5 Breast Density Classification Using M-ELBP

### 3.5.1 Overview

LBP and its variants have proven to be useful (Nanni et al., 2010; Oliver et al., 2005b) in various medical image analysis applications for extracting local texture features. A brief study on mammogram density classification is shown in Table 3.2. A detailed literature review on breast density estimation techniques are described in chapter 2.

From the above discussion, it is clear that texture has played a vital role in density segmentation/classification. Oliver et al. (2005b) combined LBP texture features and texture features extracted from grey level co-occurrence matrices to classify mammograms. Later, Chen et al. (2011) performed a comparative study on the performance of LBP, local grey-level appearance (LGA), textons and basic image features obtaining accuracies of 59%, 72%, 75% and 70%, respectively using the whole breast region from the MIAS dataset. In our study, we review the performance of various LBP variants for density classification similar to Chen et al. (2011).

The following sections analyze the impact of selecting ROI size and position for the dataset used for this experimental study. Likewise, the classification accuracy variation for different classifiers for the same set of features in a CAD process was

Table 3.2 A summary of existing approaches on breast density classification indicating authors, years of publication, techniques, experimental data for test, and classification.

| **Author** | **Features** | **Classifiers** | **Dataset** | **Number of Images** | **Result** | **Density Class** |
|---|---|---|---|---|---|---|
| (Karssemeijer, 1998) | Grey level histograms | KNN | Nijmegen | 615 | 65.0% | BIRADS I-IV |
| (Blot and Zwiggelaar, 2001) | Texture<br>GLCM | KNN | MIAS | 265 | 65.00% | Fatty, Glandular and Dense |
| (Bovis and Singh, 2002) | SGLD | ANN | DDSM | 377 | 71.40% | BIRADS I-IV |
| (Petroudi et al., 2003) | Textons based on MR8 filters | KNN | Oxford | 132 | 75.75% | 2 Class-Fatty and Dense |
| (Oliver et al., 2005b) | Morphological Features | KNN<br><br><br>Decision Tree | MIAS<br><br><br>DDSM | 270<br><br><br>300 R-MLO | 67.0%<br>73.0%<br><br>47.0% | BIRADS I-IV |
| (Muhimmah and Zwiggelaar, 2006) | Multiresolution histogram Features | DAG-SVM | MIAS | 321 | 77.57% | Fatty,Glandular and Dense |
| (Oliver et al., 2008) | Texture | KNN+SFS (sequential forward selection (SFS)) | MIAS<br><br>DDSM | 322<br><br>831 | 86.0%<br><br>77.0% | BIRADS I-IV |
| (Subashini et al., 2010) | Statistical features | SVM | MIAS | 43 | 95.55% | Fatty,Glandular and Dense |
| (Oliver et al., 2010) | Connected density clusters taking the spatial information | LDA-PCA | DDSM | 831 | 79.0% | Fatty and dense |
| (Zwiggelaar, 2010) | LGA | KNN | MIAS | 322 | 64.0% | BIRADS I-IV |
| (Liu et al., 2011) | Statistical | SVM | Tianjin Tumor Hospital | 88 | 86.40% | BIRADS I-IV |
| (Tzikopoulos et al., 2011) | Intensity based | SVM | MIAS | 322 | 85.70% | Fatty, Glandular and Dense |
| (Chen et al., 2011) | LBP<br>LGA<br>BIF<br>Textons<br>Topographic | KNN,<br>Bayesian | MIAS | 322 | 59.0%<br>72.0%<br>70.0%<br>75.0%<br>76.0% | BIRADS I-IV |
| (Qu et al., 2011) | Unknown | E-FELM<br>(Evolutionary Fuzzy Extreme Learning Machine) | MIAS | 322 | 72.6% | BIRADS I-IV |
| (Muštra et al., 2012) | GLCM | KNN+Naïve Bayesian | KBD-FER,<br>University Hospital Dubrava, Zagreb | 144 | 79.30% | 2 Class-Fatty and Dense |
| (Muštra et al., 2012) | GLCM | KNN+Naive Bayesian | MIAS | 144 | 82.0% | 2 class-Fatty, Dense |
| (He et al., 2012) | Texture | Binary model based Bayes Classifier | MIAS | 322 | 78.0% | BIRADS I-IV |
| (Rampun et al., 2017) | LTP | SVM | MIAS | 322 | 82.3% | BIRADS I-IV |
| (George et al., 2018b) | M-ELBP<br>LBP | Bayesian | MIAS | 321 | 74.0%<br>66.5% | BIRADS I-IV |

analysed. In order to estimate the efficiency of our algorithm, we compare the different LBP variants of similar topology along with traditional LBP and Local Directional Pattern (LDP). Similarly, the variation in performance based on ROI size, the strength of classifier, performance variation between ROI feature extraction and whole breast are investigated. We perform our experimental studies using the MIAS and DDSM (for ROI size analysis) database and classify the mammogram images based on MIAS-3 and BIRADS class density classification (see Section 3.6).

### 3.5.2 Methodology

We use an ROI based approach for classification for which we segment the breast tissue ROIs from the fibroglandular region of the breast where most of the dense tissue and parenchymal tissue is present. The extracted ROIs were denoised using filtering followed by feature extraction by texture descriptors. We use MIAS and DDSM datasets for the evaluation of the proposed method.

#### Pre-Processing

The mammogram images are pre-processed to extract the breast tissue region (ROI) for processing. Segmentation of the breast region is done by removing the pectoral muscles and other artefacts like orientation tags and adhesive tapes from the mammogram image using the segmentation method by Chen and Zwiggelaar (2010). The method initially separates the background region containing the annotations, labels and frames from the whole breast. Later it uses histogram thresholding, contour growing and polynomial fitting methods to remove the pectoral muscle from the breast tissue region. Segmentation of the fibroglandular disk region can improve image processing techniques for extracting the tissue characteristics of the region.

In most cases, the regions outside the fibroglandular region contain fatty tissue regardless of their BIRADS density class as most of the dense tissue patterns develop in the fibroglandular region. So extracting similar features outside the fibroglandular regions does not provide discriminative features for tissue density classification. An appropriate ROI is extracted from each mammogram image from the fibroglandular disk area of size $256 \times 256$ pixel size as shown in Figure 3.14. For extracting the fibroglandular disc region, the longest perpendicular distance from the breast boundary is considered which is usually the distance from the nipple area and then the parallel distance line from the breast boundary is considered. The intersection point is considered as the central point of the ROI region. With this central point, an ROI of size $256 \times 256$ pixels is extracted as the sample fibroglandular disc region. Noise reduction

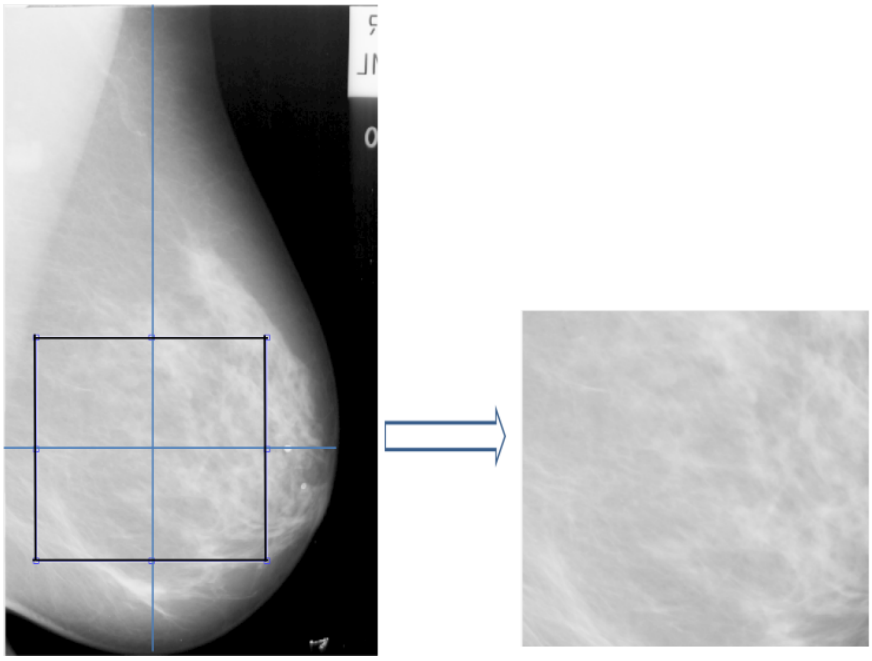


Fig. 3.14 ROI extraction from fibroglandular disk region

is then performed on the extracted ROI using a median filter of $3 \times 3$ window size (as median filter does not generate new unrealistic pixel values when the filter straddles an edge thus preserving sharp edges compared to mean filter and good at removing salt-and-pepper noise).

### 3.5.3 Feature Selection and Classification into BI-RADS Density Classes

Once the feature extraction process was completed by concatenating histogram features of all orientations, feature selection was performed to reduce feature dimensionality and to select the most prominent features for the classification. Correlation-based feature subset (CFS) selection with the best first search method was used for feature selection. Subsets of features that are highly correlated with the group class while having lower inter-correlation are preferred for attribute selection. It evaluates the individual predictive ability of each attribute/feature along with the redundancy between each of them (Hall, 1999). Later, the selected features were used by different classifiers for classifying mammographic density. A Bayesian network classifier was used for measuring the classification performance of mammograms into BI-RADS density classes in our study. In our experiments, the Bayes network uses the conditional probability estimator and a local score based structure learning algorithm. Moreover, the Bayes network is also called an independent feature model as it assumes

no relationship between the features (Witten and Frank, 2002). The graphical structures (each node connected to other nodes) are used to represent knowledge about an uncertain domain where each node represents an attribute which contains a random variable while the arch between the nodes represents the probabilistic dependencies among the corresponding random variables. The probabilistic relationship among the attributes are computed by Bayesian network algorithm based on Bayes rule (Friedman et al., 1997).The attributes in the training data are initially mapped in a form of graph topology based on the probability dependency values among the attributes which can be done by computing the probability of the random variable occurring in the training data. For example, for a random variable (A) which is dependent on another variable (B), the algorithm computes the probability of A given that B has occurred.

### CAD Result Evaluation

In order to evaluate the efficiency of proposed approach, the features are extracted and classified using various classifiers to investigate the robustness of the algorithm. The performance of CAD systems can be explained through specificity and sensitivity of the system. Sensitivity (true positives (TP) / (true positives (TP) + false negatives (FN)) is the ability of system to correctly detect the unhealthy patients (Altman and Bland, 1994). High sensitivity helps to minimize false negative detection. Similarly, specificity (true negatives (TN) / (true negatives (TN) + false positives (FP)) ) is the ability of the system to reject the healthy patients without any symptoms from the group. Likewise, other measurements used in the thesis for the efficiency of CAD systems are accuracy (performance measurement of the whole algorithm in true detection) , F-measure (harmonic mean of precision and sensitivity) , precision (fraction of relevant detection) and area under the curve ($A_z$). The area under the curve $(A_z)$ is used to evaluate the global discriminatory power of the selected classification technique. The value of $(A_z)$ ranges between 0 to 1.0. If $(A_z)$ value is high, the higher the overall performance of the classification technique would be. These CAD efficiency calculation methods will be used in our thesis for calculating classification accuracy of the developed methods.

### Experimental Results for MIAS dataset

To test the performance of the proposed M-ELBP method, the Mammographic Image Analysis Society (MIAS) database was used. Each image in the database is classified into BI-RADS classes I, II, III, IV by expert radiologists. More details on the dataset are explained in Section 2.9.

A Bayesian Network classifier was used for classification of the feature vector obtained by the application of M-ELBP operator after feature selection into BI-RADS density classes. A 10-fold cross-validation scheme was used for the classification process.

In our study, after conducting experiments with different scales, we selected appropriate parameters suitable for the extraction of mammographic texture features as: $P$=8, radius={4,7} and orientations $\theta = \{0°, 45°, 90°, 135°, 180°, 225°, 270°$ and $315°\}$ for the M-ELBP operator. With experimental results, it is suggested that the feature extraction from a single orientation is not sufficient to capture mammographic image tissue details because of the complex and multidimensional appearance of parenchymal patterns in the breast.

Table 3.3, shows the classification results of mammograms into BI-RADS classes by the Bayesian Network. The proposed method of feature extraction has achieved a classification accuracy of 75.43% with $A_z$, sensitivity, specificity, and F-measure equal to 0.92, 0.75, 0.75, and 0.75, respectively. From the classification results as shown Table 3.3, the M-ELBP operator performs better in discriminating highly dense tissue and fatty tissue compared to BI-RADS II and III classes.

Table 3.3 Confusion matrix for automatic BI-RADS classification using the proposed M-ELBP methodology on MIAS dataset.

| | | **Automatic Classification** | | | | |
|---|---|---|---|---|---|---|
| | **BI-RADS** | **I** | **II** | **III** | **IV** | **CA%** |
| **Truth Data** | **I** | 74 | 13 | 0 | 0 | 85.1 |
| | **II** | 17 | 66 | 20 | 0 | 64.1 |
| | **III** | 0 | 17 | 73 | 4 | 77.7 |
| | **IV** | 0 | 0 | 8 | 29 | 78.4 |

As per the comparison study between the conventional LBP operator (see Table 3.4), and M-ELBP operator for the selected ROI regions, the classification accuracy of LBP was 66.41% with $A_z$, specificity, sensitivity, and F-measure equal to 0.86, 0.67, 0.66, and 0.66, respectively. The classification results show that the M-ELBP is better for texture feature extraction from mammogram images for tissue classification due to the parenchymal structure of breast tissue. Likewise, a study on density classification using uniform-LBP (u-LBP) showed classification accuracy of 65.42%, slightly less than traditional LBP (see Table 3.5). The classification showed $A_z$, specificity, sensitivity, and F-measure equal to 0.87, 0.65, 0.65, and 0.65, respectively.

Table 3.4 Confusion matrix for automatic BI-RADS classification using LBP operator for MIAS dataset.

| | | **Automatic Classification** | | | | |
|---|---|---|---|---|---|---|
| | **BI-RADS** | **I** | **II** | **III** | **IV** | **CA%** |
| **Truth Data** | **I** | 68 | 19 | 0 | 0 | 78.2 |
| | **II** | 19 | 65 | 19 | 0 | 63.1 |
| | **III** | 3 | 28 | 54 | 9 | 57.4 |
| | **IV** | 0 | 0 | 11 | 26 | 70.3 |

Table 3.5 Confusion matrix for automatic BI-RADS classification using u-LBP operator for MIAS dataset.

| | | **Automatic Classification** | | | | |
|---|---|---|---|---|---|---|
| | **BI-RADS** | **I** | **II** | **III** | **IV** | **CA%** |
| **Truth Data** | **I** | 70 | 13 | 3 | 1 | 80.5 |
| | **II** | 15 | 59 | 27 | 2 | 57.3 |
| | **III** | 0 | 32 | 54 | 8 | 57.4 |
| | **IV** | 2 | 2 | 6 | 27 | 73.0 |

From the M-ELBP classification results, we included additional features to improve the classification results. As an initial step, we included image features like maximum intensity, minimum intensity, entropy and standard deviation with the M-ELBP feature vector. Table 3.6 shows the classification results by Bayesian Network when additional image features were added to the M-ELBP texture descriptor. We found from the corresponding experiment that the overall classification accuracy (CA%) remained closed to 75.4% with $A_z$, specificity, sensitivity, and F-measure as 0.92, 0.75, 0.75, and 0.75, respectively. While looking into class-wise classification, BI-RADS I and II were classified better with more than 80.0%. This shows that with additional features, class-wise classification can be achieved.

Similarly, when the effect of applying a mean filter before an ELBP operator is compared to the proposed M-ELBP variant, we found that the overall classification accuracy of ELBP to be 67.00% as shown in Table 3.7. The $A_z$, specificity, sensitivity, and F-measure and was found to be 0.88, 0.66, 0.67, and 0.66, respectively.

In order to evaluate the effect of mean-filter before an ELBP operator to ELBP operator, a mean filter was applied before performing ELBP feature extraction. The

Table 3.6 Confusion matrix for automatic BI-RADS classification using the proposed M-ELBP methodology with image intensity features for MIAS dataset.

| | | Automatic Classification | | | | |
|---|---|---|---|---|---|---|
| | **BI-RADS** | **I** | **II** | **III** | **IV** | **CA%** |
| **Truth Data** | **I** | 74 | 13 | 0 | 0 | 85.1 |
| | **II** | 16 | 70 | 17 | 0 | 68.0 |
| | **III** | 0 | 21 | 68 | 5 | 72.3 |
| | **IV** | 0 | 0 | 7 | 30 | 81.1 |

Table 3.7 Confusion matrix for automatic BI-RADS classification using ELBP for MIAS dataset.

| | | Automatic Classification | | | | |
|---|---|---|---|---|---|---|
| | **BI-RADS** | **I** | **II** | **III** | **IV** | **CA%** |
| **Truth Data** | **I** | 69 | 18 | 0 | 0 | 79.3 |
| | **II** | 20 | 57 | 26 | 0 | 55.3 |
| | **III** | 0 | 24 | 59 | 11 | 62.8 |
| | **IV** | 0 | 0 | 7 | 30 | 81.1 |

Table 3.8 shows a classification accuracy of 69.73%. $A_z$, specificity, sensitivity, and F-measure and was found to be 0.89, 0.69, 0.69, and 0.69, respectively. Table 3.8 shows the detailed confusion matrix for the mean filter before ELBP operator using the Bayesian classifier. The confusion matrix shows that the features extracted by applying a mean filter before an ELBP operator classifies the BIRADS I and IV much better than BIRADS III and IV. The study shows that the classification performance is better for M-ELBP operator compared to ELBP and mean filter applied before an ELBP operator.

In order to evaluate the efficiency of the proposed method over the traditional texture extraction method like Gray-Level Co-Occurrence Matrix (GLCM), we extracted the Haralick features from the same set of data used for evaluating the binary patterns. It was found that the proposed method performed better than the GLCM for the selected data. The standard method classified the dataset into BI-RADS classes and obtained a classification accuracy of 55.80% with $A_z$, specificity, sensitivity, and F-measure of 0.77, 0.59, 0.56, and 0.55, respectively using Bayesian classifier. More

Table 3.8 Confusion matrix for automatic BI-RADS classification using mean filter and ELBP for MIAS dataset.

| | | **Automatic Classification** | | | | |
|---|---|---|---|---|---|---|
| | **BI-RADS** | **I** | **II** | **III** | **IV** | **CA%** |
| **Truth Data** | **I** | 75 | 12 | 0 | 0 | 86.2 |
| | **II** | 16 | 58 | 26 | 3 | 56.3 |
| | **III** | 0 | 23 | 60 | 11 | 63.8 |
| | **IV** | 0 | 0 | 6 | 31 | 83.8 |

details of the classification are shown in Table 3.9. It shows that the proposed method performs better than the standard texture analysis methods like GLCM and LBP.

Table 3.9 Confusion matrix for automatic BI-RADS classification using GLCM features for MIAS dataset.

| | | **Automatic Classification** | | | | |
|---|---|---|---|---|---|---|
| | **BI-RADS** | **I** | **II** | **III** | **IV** | **CA%** |
| **Truth Data** | **I** | 54 | 29 | 1 | 3 | 62.1 |
| | **II** | 14 | 75 | 11 | 3 | 72.8 |
| | **III** | 5 | 34 | 33 | 22 | 35.1 |
| | **IV** | 0 | 8 | 12 | 17 | 45.9 |

The experimental results suggest that the M-ELBP are able to extract robust and intrinsic features from mammographic images especially with dense tissues. Moreover, it is able to capture detailed micro patterns from the mammogram giving more detailed texture descriptor features for classification. The use of image regions from the fibroglandular disc region reduced the chance of extracting less discriminating features for classification and it also reduced the computational complexity. With the choice of extracting features from 8 different orientations, we are able to extract more micro-structure features from the mammogram considering the complexity and multi dimensional structure of parenchymal pattern in the breast.

Similarly, we performed the classification of the whole breast region in-order to find the difference in classification accuracy compared to the extracted ROI classification. The whole breast region classification for BI-RADS class obtained an accuracy of 57.0% for the MIAS dataset using the Bayesian Network. The classification obtained $A_z$, specificity, sensitivity, F-measure and of 0.81, 0.57, 0.57, and 0.56, respectively.

Detailed confusion matrix for the classification is shown in Table 3.10. This could be because of the irrelevant or repeated features in the outer regions of breast region.

Table 3.10 Confusion matrix for automatic BI-RADS classification using M-ELBP for whole breast region for MIAS dataset.

| | | **Automatic Classification** | | | | |
|---|---|---|---|---|---|---|
| | **BI-RADS** | **I** | **II** | **III** | **IV** | **CA%** |
| **Truth Data** | **I** | 72 | 15 | 0 | 0 | 82.8 |
| | **II** | 19 | 47 | 33 | 4 | 45.6 |
| | **III** | 2 | 30 | 37 | 25 | 39.4 |
| | **IV** | 0 | 1 | 9 | 27 | 73.0 |

Table 3.11 provides a better understanding of the efficiency of the proposed method over the traditional methods like LBP, u-LBP, GLCM features, ELBP, applying a mean filter before ELBP, u-ELBP and M-ELBP. The study shows that M-ELBP was performing better than other methods in the list and was better at classifying the BIRADS I and IV.

Table 3.11 Comparison of CA% and $A_z$ for LBP variants with proposed M-ELBP operator for the classification of MIAS dataset into BIRADS I-IV classes.

| **Feature Extractor** | **CA%** | $A_z$ |
|---|---|---|
| GLCM | 55.80 | 0.77 |
| LBP | 66.41 | 0.86 |
| u-LBP | 65.42 | 0.87 |
| ELBP | 67.00 | 0.88 |
| mean filter+ELBP | 69.73 | 0.89 |
| u-ELBP | 68.34 | 0.87 |
| **M-ELBP** | **75.43** | **0.92** |

**Discussion**

The experimental study on the MIAS dataset showed that the M-ELBP operator was able to classify the mammographic images better than the traditional LBP and standard techniques like GLCM texture features. Likewise, the classification performed better on extracted ROI's from the fibroglandular region than using the whole breast region.

Similarly, the new operator is more resistant to noise in the image as it is not comparing the pixels directly as in classic LBP. M-ELBP was found to extract more features from different directions as it is not rotation invariant and it can be used to obtain features from different directions and varying scales giving it a multi-scale feature extraction ability.

The confusion matrix shows that the proposed methods performs better for both BI-RADS I and IV. The studies show that BI-RADS II and III have similar mammographic structure and is difficult to categories them resulting in inter and intra-classification variations by radiologists. On the other hand, BI-RADS I and IV have comparatively different structure resulting in better classification Muhimmah et al. (2006); Redondo et al. (2012). It is to be noted that most of the research works in the literature uses MIAS 3-class density classification for density classification as the ground truth is available with the MIAS dataset. This can reduce the intra-observable error rate when classified by different radiologists. Therefore, we use the 3-class density classification for further experimental studies.

## 3.6 Tissue Based 3-Class Density Classification-MIAS

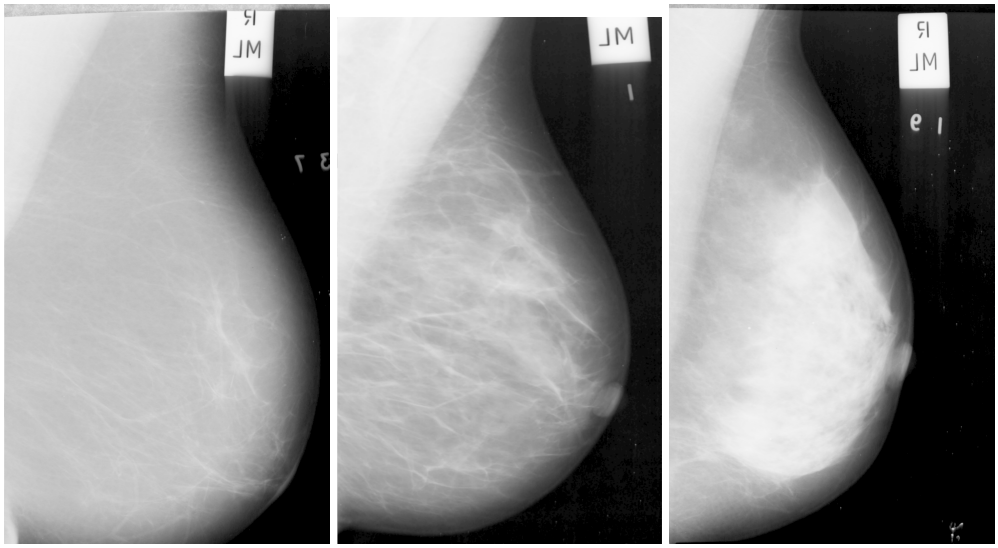
From the experimental study, it can be observed that the M-ELBP operator could classify better than traditional LBP. However, the effect of descriptor size, the ROI size and location, the effect of selecting ROIs and whole breast, and the classifier effect on classification should also be investigated. Similarly, the role of directional filters in mammographic feature extraction needs to be included as well. To this end, a texture directional filter, and a Local Directional Filter (LDP) has been included for further investigation.

### 3.6.1 A Comparative Study with other LBP variants

To investigate the feature extraction capability of different LBP variants, we focus on traditional LBP, ELBP, u-ELBP, M-ELBP and LDP.

In our study, for evaluation purposes, we use the MIAS dataset (Suckling et al., 1994). The MIAS dataset provides a three class density classification (F,G,D) as ground truth and is publically available, we use it in our comparative study as well. This helps in controlling the inter and intra-validation variation effects while classifying the mammographic images by different radiologists (Kauppi et al., 2009; Warfield et al., 2006). More details on the dataset can be found in Section 2.9.1. Examples of

mammographic images with different tissue type from the MIAS dataset are shown in Figure 3.15.



(a) (b) (c)

Fig. 3.15 Example of mammographic tissue types: (a) fatty, (b) glandular, and (c) dense.

To study the importance of directional texture feature descriptors for extracting mammographic parenchymal features, we compute the classification accuracy of the M-ELBP and LDP developed by Jabid et al. (2010). The LDP produces an eight-bit binary pattern similar to LBP for a neighborhood of eight. Unlike LBP, the LDP patterns are computed by comparing the relative edge response values of each pixel at different orientations. To get the edge responses, the eight directional edge response value of a particular pixel is calculated using Kirsch masks ($(M_0 - M_7)$, also known as the Kirsch compass kernel which is a non-linear edge detector to find the maximum edge strength in eight predetermined compass directions with a 45-degree increments). The edge responses by masks are not equally important in all directions ($m_0, m_1 .... m_7$) as the corners and edges show high responses for particular directions. So the top $k$ values from $|m_j|$ were selected and defined as '1' with the other $(8-k)$ bits as '0'. The detailed representation of LDP pattern generation and Kirish mask is shown in Figure 3.16 and Figure 3.17

For the comparison of classification by different LBP variants based on MIAS tissue-based classification, we performed LBP, ELBP, Uniform-ELBP (u-ELBP), M-ELBP and LDP on ROIs of size $256 \times 256$ from the fibroglandular region with neighborhood size equal to eight. While LBP extracted features only from a circular neighborhood, the ELBP was able to extract more structural and spatial features

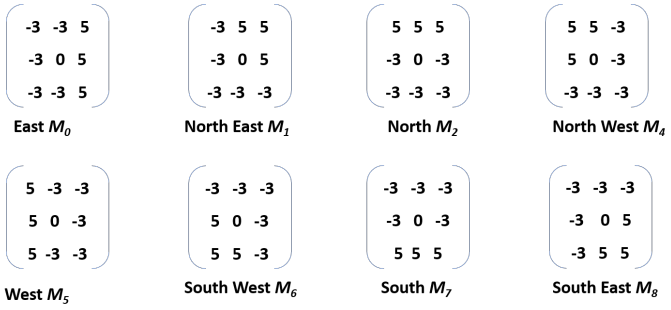



Fig. 3.16 Kirsch Edge Mask in all eight directions.

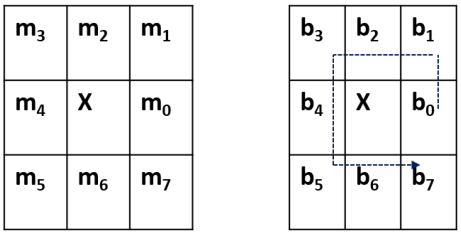



Fig. 3.17 LDP Pattern generation:-Left column: Eight directional edge response positions, Right column: LDP binary bit positions.

at different orientations extracting multi-dimensional micro-pattern features of the breast tissue. To incorporate the intensity features along with the textural features into the histogram, we used M-ELBP. Similarly, the effect of uniform patterns in tissue classification was estimated using u-ELBP (Nguyen and Caplier, 2012). u-ELBP can reduce the length of the feature vector from 256 to 59 for an 8- bit binary pattern. The uniform pattern concept was developed from the fact that certain binary patterns occur commonly in texture images (A detailed description of the uniform pattern is explained in Section 3.3.1). To keep the consistency of classification accuracy due to the classifier effect, we used the Bayesian Network for every descriptor with 10-fold cross-validation.

Table 3.12 Confusion matrix for automatic tissue classification using LBP descriptor for ROI size $256 \times 256$.

| | **Automatic Classification** | | | | |
|---|---|---|---|---|---|
| **Truth Data** | **Tissue Type** | **Fatty** | **Glandular** | **Dense** | **CA%** |
| | **Fatty** | 86 | 19 | 1 | 81.1 |
| | **Glandular** | 17 | 72 | 15 | 69.2 |
| | **Dense** | 3 | 37 | 71 | 64.0 |

Table 3.12 shows the classification results on the ROIs from the MIAS database using LBP. The approach gave a classification accuracy equal to 69.44%. The $A_z$, specificity, sensitivity and F-measure found to be 0.87, 0.71, 0.70, 0.70, respectively. Since the LBP operator considered a circular pattern of the neighborhood, it could not capture the directional features of the breast tissue due to its rotation invariant feature.

Table 3.13 Confusion matrix for automatic tissue classification using the ELBP descriptor for ROI size $256 \times 256$ and $R = \{4,7\}$ and $P = 8$.

| | **Automatic Classification** | | | | |
|---|---|---|---|---|---|
| **Truth Data** | **Tissue Type** | **Fatty** | **Glandular** | **Dense** | **CA%** |
| | **Fatty** | 91 | 14 | 1 | 85.8 |
| | **Glandular** | 10 | 71 | 23 | 68.3 |
| | **Dense** | 1 | 24 | 86 | 77.5 |

Table 3.13 shows that the classification accuracy for the ELBP descriptor has improved to 75.41% compared to LBP, showing that ELBP can perform better by extracting additional multidimensional features from different directions. ELBP descriptor was able to obtain an $A_z$, specificity, sensitivity and F-measure of 0.90, 0.72, 0.70, 0.70 respectively. In order to study the effect of uniform patterns in the classification of density, we performed u-ELBP giving an accuracy of 72.86% as shown in Table 3.14 with $A_z$, specificity, sensitivity and F-measure to be 0.89, 0.73, 0.73 and 0.73 respectively. The study shows that ELBP descriptors performs better than LBP, especially in classifying dense and fatty tissue types. This is due to extraction of more intrinsic features from the ROIs.

Table 3.14 Confusion matrix for automatic tissue classification using the u-ELBP descriptor for ROI size $256 \times 256$ and $R = \{4,7\}$ and $P = 8$.

| | **Automatic Classification** | | | | |
|---|---|---|---|---|---|
| **Truth Data** | **Tissue Type** | **Fatty** | **Glandular** | **Dense** | **CA%** |
| | **Fatty** | 86 | 19 | 1 | 81.1 |
| | **Glandular** | 11 | 69 | 24 | 66.3 |
| | **Dense** | 2 | 27 | 82 | 73.9 |

Furthermore, using the M-ELPB operator, a classification accuracy of 77.4% as shown in Table 3.15 with an $A_z$, specificity, sensitivity and F-measure to be 0.92, 0.79, 0.78 and 0.78, respectively is obtained. Table 3.16 shows that the classification accuracy for this ELBP variant (M-ELBP) improved to 78.8% when including more intensity features to the texture pattern. The $A_z$, specificity, sensitivity and F-measure score found to be 0.92, 0.80, 0.79 and 0.79, respectively. The study shows that improving the feature extraction methods could help in improving classification.

Table 3.15 Confusion matrix for automatic tissue classification using the M-ELBP descriptor for ROI size $256 \times 256$ and $R = \{4,7\}$ and $P = 8$.

| | **Automatic Classification** | | | | |
|---|---|---|---|---|---|
| **Truth Data** | **Tissue Type** | **Fatty** | **Glandular** | **Dense** | **CA%** |
| | **Fatty** | 93 | 13 | 0 | 87.7 |
| | **Glandular** | 12 | 73 | 19 | 70.2 |
| | **Dense** | 1 | 25 | 85 | 76.6 |

Table 3.16 Confusion matrix for automatic tissue classification using the M-ELBP descriptor for ROI size $256 \times 256$ and $R = \{4,7\}$ and $P = 8$ and image intensity features.

| | **Automatic Classification** | | | | |
|---|---|---|---|---|---|
| **Truth Data** | **Tissue Type** | **Fatty** | **Glandular** | **Dense** | **CA%** |
| | **Fatty** | 90 | 16 | 0 | 84.9 |
| | **Glandular** | 11 | 77 | 16 | 74.0 |
| | **Dense** | 1 | 24 | 86 | 77.5 |

Similarly, to analyze the role of directional texture descriptor for mammographic density estimation, we used the LDP (Jabid et al., 2010) operator on the same set of ROIs with the Bayesian Network. The classification results in Table 3.17 shows an accuracy of 75.0%, which is lower than the M-ELBP descriptor for the Bayesian Network classifier but it shows the strength of pixel directional response in feature extraction. The $A_z$, specificity, sensitivity and F-measure score found to be 0.91, 0.76, 0.75 and 0.76 respectively.

Table 3.17 Confusion matrix for automatic tissue classification using the LDP descriptor for ROI size $256 \times 256$.

| | **Automatic Classification** | | | | |
|---|---|---|---|---|---|
| **Truth Data** | **Tissue Type** | **Fatty** | **Glandular** | **Dense** | **CA%** |
| | **Fatty** | 91 | 15 | 0 | 85.8 |
| | **Glandular** | 8 | 71 | 25 | 68.3 |
| | **Dense** | 2 | 30 | 79 | 71.2 |

When considering Tables 3.12-3.17, it seems that there is a trend for over-estimation, which could be due to the wide variation in glandular tissue appearance within the dataset and the ROI selection process. From the classification results for M-ELBP (see Table 3.15), fatty and dense tissue were classified better. This can be attributabed to the incorporation of texture features along with the mean intensity values of the neighborhood making it more robust to the tissue pattern and intensity variations. In addition, it should be noted that the LBP based approach performs worse for the dense

class (which is probably caused by the rotational invariant aspects), whilst both ELBP and u-ELBP perform less on Glandular cases (which could again be caused by the wide range of variation in the Glandular class and the effects of noise). We see a more detailed analysis of these aspects as future work, which could include an in-depth analysis of the feature space.

In order to analyze the significance of the test for different texture descriptors (Stąpor, 2017), a 10 run 10 fold cross-validation for all the texture descriptors is performed. Table 3.18 shows that the Bayesian Network yielded an accuracy of $69.44 \pm 0.92$, $75.41 \pm 1.05$, $72.86 \pm 1.06$, $77.38 \pm 1.06$ and $74.92 \pm 0.67$ for LBP, ELBP, u-ELBP, M-ELBP and LDP respectively. Table 3.18 also indicates if the differences were significant (at the 0.05 level).

### 3.6.2 Effect of Descriptor Size on Classification

In order to analyse the effect of M-ELBP descriptor's size on classification, the size of descriptor parameters and perform the experiment on the MIAS dataset is varied as an initial study . The M-ELBP operator with varying sizes ranging from $R1 = \{2,4,6,12\}$ and $R2 = \{5,7,9,15\}$ were applied for the ROI size equal to $256 \times 256$ pixels. To compute the stability of the approach, we used a 10 run 10 fold cross-validation scheme using a Bayesian Network.

#### Experimental Results and Discussion

Figure 3.18 shows the classification accuracy ($CA\%$) as $77.82 \pm 0.68$, $77.66 \pm 1.01$, $78.01 \pm 0.57$ and $76.85 \pm 0.73$. The figure represents the classification accuracy obtained by the Bayesian Network classifier for each run ($N$= run number) for a 10-fold cross-validation scheme. Similarly from Figure 3.19, the area under the ROC curve ($A_z$) shows similar results as $0.9204 \pm 0.0010$, $0.9181 \pm 0.0011$, $0.9205 \pm 0.0015$ and $0.9121 \pm 0.0022$ for the M-ELBP descriptor of size $R1 = \{2,4,7,12\}$, $R2 = \{5,7,9,15\}$ with a neighborhood of eight pixels, respectively. It is to be noted that if the descriptor size is made too large, the descriptor will not be able to extract texture features near to the central pixel, e.g. for $\{R1,R2\}$=$\{12,15\}$. Moreover, the $CA\%$ is slightly degraded with larger descriptor size, however, on the other hand ($A_z$) shows little variation.

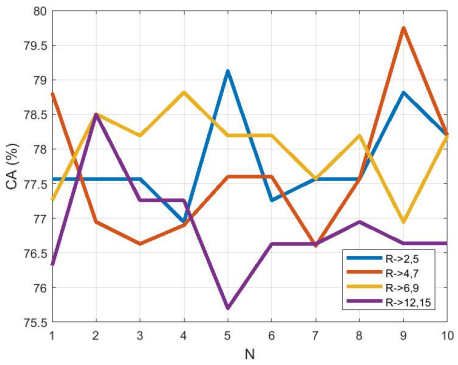



Fig. 3.18 Classification accuracy as a function of descriptor size (N indicates a distinct run).

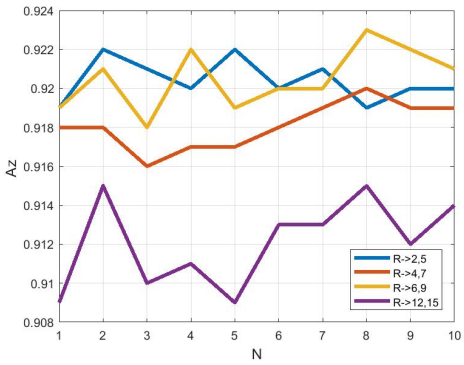



Fig. 3.19 The area under ROC curve $A_z$ as a function of descriptor size (N indicates a distinct run).

### 3.6.3 Role of Classifier in Classification Accuracy

Involvement of machine learning algorithm in developing CAD systems for medical image classification is a promising field of research (Nahid and Kong, 2017). So, in our experimental study, we investigate the effect of different classifiers on breast density classification using LBP variants. For consistency the descriptor size and ROI size, we choose ROIs of size $256 \times 256$ pixels from the fibroglandular region and a descriptor neighbourhood of size $P = 8$.

The classification accuracy illustrated by Table 3.18 shows the variation in classification accuracy for a 10 run 10 fold cross-validation for different classifiers showing the importance of feature selection and classifier effect on density classification. The classifiers chosen for our experimental study are Bayesian Network, K-Nearest neighbor classifier, Support Vector Machine, and Random Forest (Bishop, 1995; Duda and Hart, 1973). To test the statistical significance of classification accuracy, we have performed a paired t-test for a significance level of $p = 0.05$ between descriptors and classifiers (see Table 3.19) taking best classification result by the respective classifier for each descriptor as the base line (i.e. across rows/columns). A paired t-test (also known as a dependent or correlated t-test) is a statistical test that compares the averages/means and standard deviations of two related groups to determine if there is a significant difference between the two groups. A significant difference occurs when the differences between groups are unlikely to be due to sampling error or by chance. Similarly, the groups can be related by being the same group of people, the same item, or being subjected to the same conditions.

Table 3.18 shows the classification results of LBP variants for each classifier while Table 3.19 shows the $p$-values obtained for the t-test performed (i.e. across

Table 3.18 CA% results by various classifiers for LBP variants for ROI size $256 \times 256$ pixels for 10 run 10 fold cross-validation.

| Classifier | **LBP** | **ELBP** | **u-ELBP** | **M-ELBP** | **LDP** |
|---|---|---|---|---|---|
| **Bayesian Network** | $69.44 \pm 0.92$ | $75.41 \pm 1.05$ | $72.86 \pm 1.06$ | $\mathbf{77.38 \pm 1.06}$ | $74.92 \pm 0.67$ |
| **KNN** | $69.34 \pm 1.09$ | $70.43 \pm 1.03$ | $71.46 \pm 1.41$ | $75.46 \pm 0.78$ | $\mathbf{75.95 \pm 0.96}$ |
| **SVM** | $67.07 \pm 0.75$ | $73.11 \pm 0.76$ | $73.29 \pm 0.64$ | $\mathbf{74.42 \pm 0.92}$ | $74.17 \pm 0.52$ |
| **Random Forest** | $69.19 \pm 1.31$ | $73.26 \pm 1.08$ | $73.04 \pm 0.97$ | $73.20 \pm 1.36$ | $\mathbf{75.91 \pm 1.01}$ |

Table 3.19 The $p$ values between the best CA% value of each descriptor against the other classifiers for same set (the $p$ value set as 0.05). Bold values indicate statistically insignificant.

| Classifiers/Descriptors | **LBP** | **ELBP** | **u-ELBP** | **M-ELBP** | **LDP** |
|---|---|---|---|---|---|
| **Bayesian Network** | $< 0.0001$ | $< 0.0006$ | $< 0.0001$ | $-$ | $< 0.0001$ |
| **KNN** | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ | **=0.2263** | $-$ |
| **SVM** | $< 0.0001$ | $= 0.0027$ | $= 0.0051$ | $-$ | **=0.4641** |
| **Random Forest** | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ | $-$ |

rows/columns). The significance value set for the experiment was 0.05. A significance level of 0.05 indicates a 5% risk of concluding that a difference exists when there is no actual difference. The highest classification value obtained for each classifier was tested against other descriptors. Table 3.19 depicts that most of the test were statistically significant as the $p$ values were less than or equal to the significance level. The values represented in bold indicate that they are statistically insignificant as the tested $p$ values were above the significance level.

Table 3.20 Area under ROC ($A_z$) classification results for LBP variants using ROI size $256 \times 256$ pixels for a 10 run 10 fold cross validation.

| Classifier | **LBP** | **ELBP** | **Uniform ELBP** | **M-ELBP** | **LDP** |
|---|---|---|---|---|---|
| **Bayesian Network** | $0.89 \pm 0.006$ | $0.90 \pm 0.001$ | $0.89 \pm 0.003$ | $\mathbf{0.92 \pm 0.000}$ | $0.89 \pm 0.004$ |
| **KNN** | $0.86 \pm 0.005$ | $0.85 \pm 0.003$ | $0.87 \pm 0.005$ | $0.89 \pm 0.005$ | $\mathbf{0.89 \pm 0.002}$ |
| **SVM** | $0.78 \pm 0.004$ | $0.83 \pm 0.005$ | $0.84 \pm 0.003$ | $\mathbf{0.85 \pm 0.005}$ | $0.84 \pm 0.002$ |
| **Random Forest** | $0.86 \pm 0.004$ | $0.89 \pm 0.004$ | $0.89 \pm 0.004$ | $0.89 \pm 0.004$ | $\mathbf{0.90 \pm 0.003}$ |

Similarly, the area under the ROC curve ($A_z$) (Bradley, 1997) was calculated to compare the efficiency of classifiers for density estimation. Table 3.20 shows results using 10 run 10 fold cross-validation for $A_z$. The $p$ values in Table 3.21 (i.e, across

Table 3.21 The $p$ values between the best ROC value of each descriptor against the other classifiers for the same set ($p$ value set as 0.05). Bold values indicate statistically insignificant.

| Classifiers/Descriptors | **LBP** | **ELBP** | **u-ELBP** | **M-ELBP** | **LDP** |
|---|---|---|---|---|---|
| **Bayesian Network** | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ | – | $< 0.0001$ |
| **KNN** | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ | **=0.9169** | – |
| **SVM** | $< 0.0001$ | $< 0.0001$ | $< 0.0002$ | – | **=0.4328** |
| **Random Forest** | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ | **=0.0830** | – |

rows/columns) indicate that most of the test were statistically significant while cells represented in bold are statistically insignificant for the $A_z$ results shown in Table 3.20 as they are greater than the significance value 0.05.

On comparing the classification accuracy and $A_z$ results from Tables 3.18 and 3.20, LDP and M-ELBP show a balanced result for classification accuracy by classifiers showing the potential of directional filters for mammogram density classification.

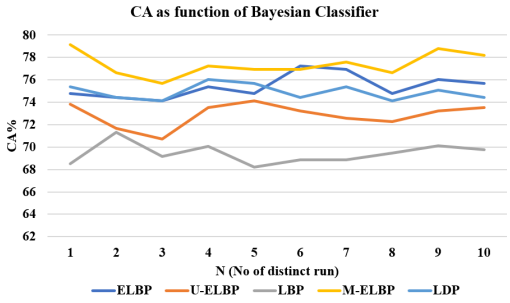



Fig. 3.20 Classification accuracy for Bayesian network for LBP variants.

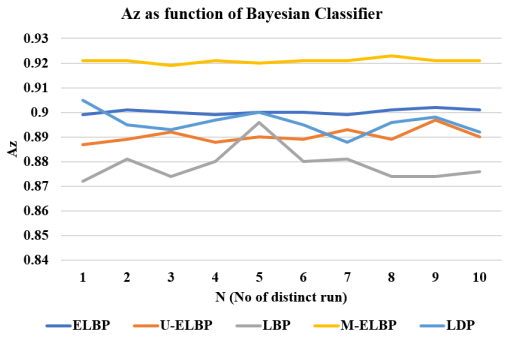



Fig. 3.21 The area under ROC curve $A_z$ for Bayesian network for LBP variants.

Similarly, Figures 3.20 and 3.21 represent the $A_Z$ values for LBP variants using the Bayesian classifier. The results show that M-ELBP performed better compared to other variants for the Bayesian Network. From the experimental studies, the M-ELBP performed better in classifying the mammographic breast tissue showing the importance of directional filters in feature extraction for complex texture images like mammograms.

### Experimental Results and Discussion

The experimental study for classifiers on the same feature set shows slight variations in classification accuracy as shown in Figures 3.22-3.26. This shows that different

classifiers have varying techniques in analysing the features and classifying them. So, choosing a better classifier with appropriate learning algorithms can improve the accuracy of the CAD system. The unpaired t-test with a p=0.05 value was performed on the best classification result for each classifier to the descriptors CA% and ROC value to estimate the statistical significance of the results (see Table 3.19 and Table 3.21).

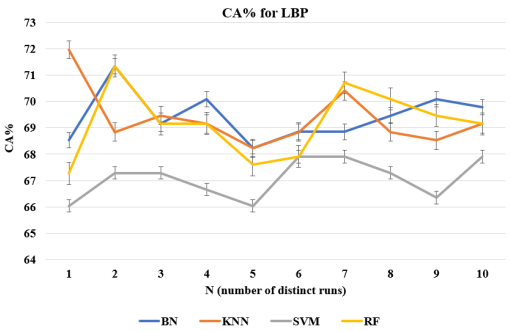



Fig. 3.22 LBP CA% as a function of classifiers.

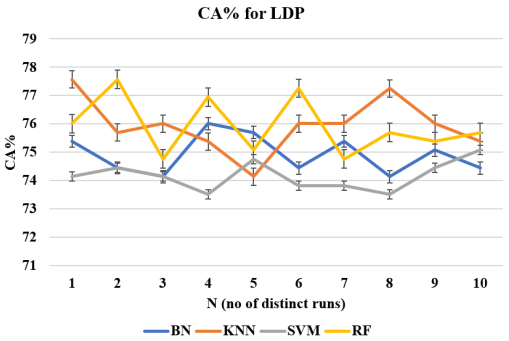



Fig. 3.23 LDP CA% as a function of classifiers.

### 3.6.4 Effect of Selecting ROI Size

A comparison study was performed between pre-processed whole mammogram and ROIs of size $256 \times 256$ to investigate the effect of selecting ROIs for classification using the MIAS dataset in Section 3.5.3. The study showed that classification on selected ROIs performed better than the whole image. To investigate the effect of ROI sizes in breast density classification using the proposed M-ELBP operator and the available datasets (MIAS and DDSM), varying sizes of ROIs were extracted from both datasets. The features extracted from both datasets were classified using BI-RADS classes and the efficiency was analysed using CA% and $A_z$. The M-ELBP descriptor of size $R = \{4,7\}$ and $P = 8$ was used for feature extraction and the Bayesian Network was used for the classification analysis using a 10 run 10 fold cross validation scheme.

As an initial study, varying sizes of ROIs were extracted from the MIAS dataset ranging from $256 \times 256$ pixels, $128 \times 128$ pixels, $64 \times 64$ pixels and $32 \times 32$ randomly from the fibroglandular region. Increasing the size of ROIs to $512 \times 512$ resulted in extracting regions outside breast area for the selected dataset. Therefore, we chose ROIs from $256 \times 56$ for the experimental study. As discussed in Section 3.5.3, the features were extracted using the M-ELBP descriptor and classified using the Bayesian Network classifier.

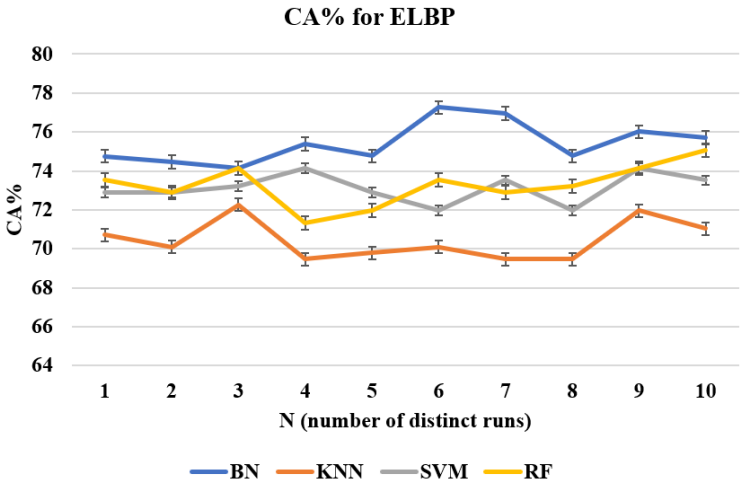



Fig. 3.24 ELBP CA% as a function of classifiers.

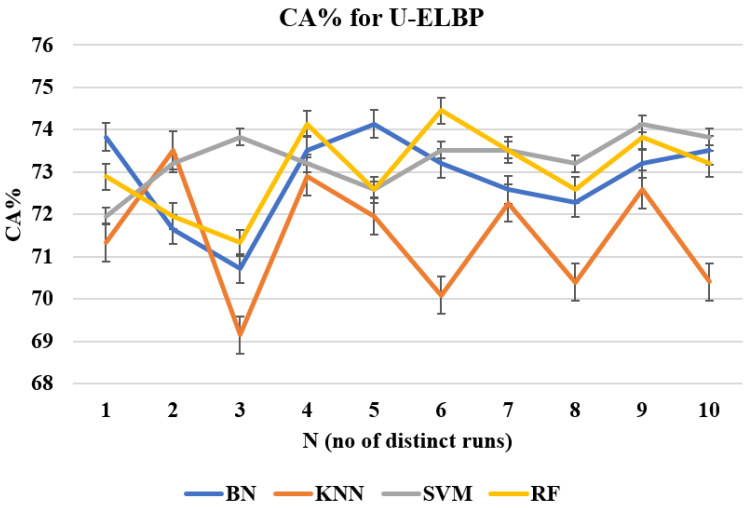



Fig. 3.25 u-ELBP CA% as a function of classifiers.

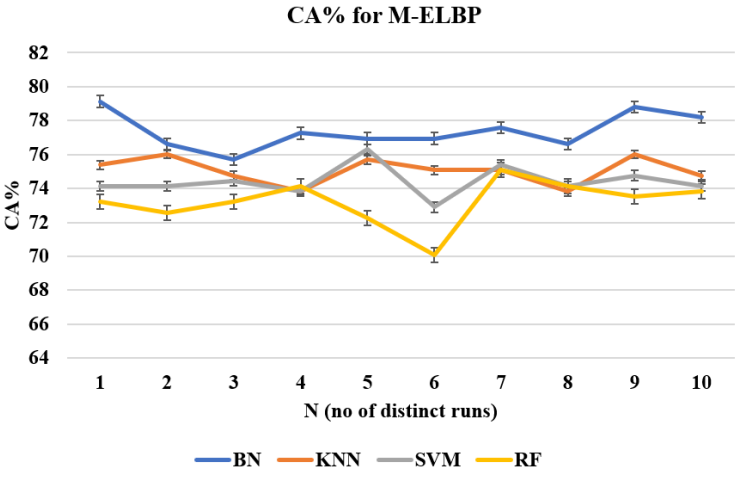



Fig. 3.26 M-ELBP CA% as a function of classifiers.

Table 3.22 CA% and $A_z$ as a function of varying patch size for MIAS dataset BI-RADS density classification

| | **Patch size** | **CA%** | $A_z$ |
|---|---|---|---|
| **ROI size** | **256x256** | **74.04 ± 0.750** | **0.91 ± 0.001** |
| | **128x128** | 66.93 ± 0.853 | 0.87 ± 0.003 |
| | **64x64** | 59.18 ± 1.377 | 0.80 ± 0.004 |
| | **32x32** | 51.88 ± 1.550 | 0.75 ± 0.004 |

Table 3.22 shows the CA% and $A_z$ values for BI-RADS classification for varying sizes of ROIs from MIAS dataset. Figures 3.27 and 3.28 shows the graphical representation of CA% and $A_z$ for different patches at each run. It is to be noted that the best classification was found to be for the patch $256 \times 256$ pixels and decreases with the patch size.

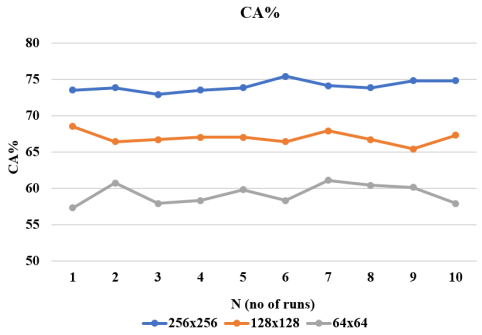



Fig. 3.27 CA% as a function of ROI sizes for MIAS dataset.

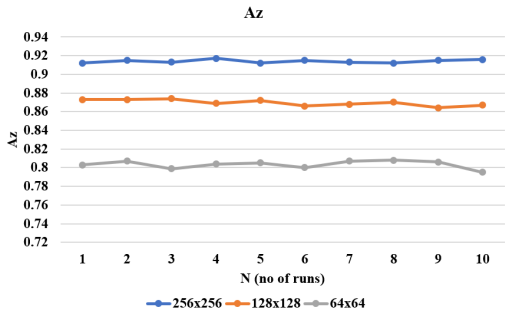



Fig. 3.28 $A_z$ as a function of ROI sizes for MIAS dataset.

To get further insight, the same set of experiment were performed on the DDSM dataset and classified the images into BI-RADS classes. A set of 450 mammogram images from DDSM dataset, (57, 179, 128 and 86 images were BI-RADS I-IV respectively) was used. More details on DDSM dataset are provided in Section 2.9. Similar to MIAS, we extracted ROI patches ranging from $256 \times 256$ pixels, $128 \times 128$ pixels and $64 \times 64$ pixels. The M-ELBP descriptor of size $R = \{4,7\}$ and $P = 8$ was used for feature extraction and the Bayesian Network was used for the classification analysis using a 10 run 10 fold cross evaluation scheme to match the experiment on the MIAS dataset. Though the classification results were poor compared to MIAS, the overall results for both datasets match closely. For both cases, the classification accuracy decreases with the patch size. The CA% and $A_z$ for the patches $256 \times 256$ pixels, $128 \times 128$ pixels and $64 \times 64$ pixels was found to be $47.65 \pm 1.06$, $45.82 \pm 0.50$, $38.33 \pm 0.72$ and $0.71 \pm 0.01$, $0.62 \pm 0.01$, $0.59 \pm 0.01$ respectively. A graphical illustration is shown in Figures 3.29 and 3.30. It shows that the ROI patch $256 \times 256$

pixels performed better compared to other ROIs. The confusion matrix shown in Table 3.23 describes an overview of the BI-RADS classification for DDSM with CA%, $A_z$, specificity, sensitivity and F-measure equal to 47.65%, 0.71, 0.47, 0.48 and 0.47, respectively.

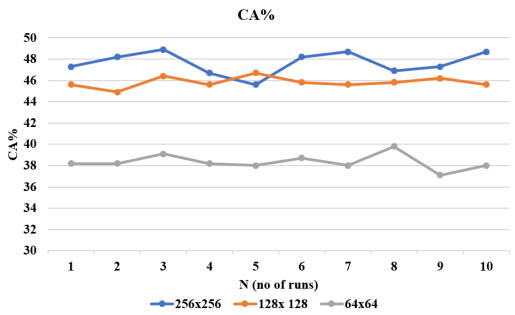



Fig. 3.29 CA% as a function of ROI sizes for DDSM dataset.

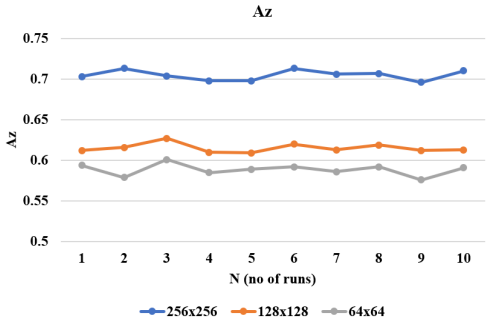



Fig. 3.30 $A_z$ as a function of ROI sizes for DDSM dataset.

Table 3.23 Confusion matrix for automatic BI-RADS classification using M-ELBP for $256 \times 256$ region for DDSM dataset.

| | | **Automatic Classification** | | | | |
|---|---|---|---|---|---|---|
| | **BI-RADS** | **I** | **II** | **III** | **IV** | **CA%** |
| **Truth Data** | **I** | 16 | 23 | 13 | 5 | 28.1 |
| | **II** | 20 | 102 | 44 | 13 | 57 |
| | **III** | 1 | 44 | 59 | 24 | 46.1 |
| | **IV** | 3 | 12 | 28 | 43 | 50.0 |

The study showed that if the ROI size is small, classification can be affected due to minimum dominant features. Similarly, if the ROI size is too large or include irrelevant regions like tissue near the pectoral muscle or the breast boundary regions, it can result in extracting few discriminant features. In the case of choosing the whole breast, the classification can be affected by irrelevant data resulting in a poor classification accuracy as it extracts multi-class texture information making it difficult for discriminating the prominent features for each density class. So the study showed that choosing the correct ROI size and location can lead to improved density classification in line with similar studies in the field (Li et al., 2004; Sharma, 2017). The experimental study showed that selecting the size of the ROIs around $256 \times 256$ pixels taken from fibroglandular disk region is most appropriate for improved classification. This showed a similar result to the study performed by Li et al. (2004) where ROIs were selected

at different distances from the breast nipple. In addition, a similar study performed by Rampun et al. (2018a) showed the importance of choosing the ROIs from the fibroglandular region for density classification. Further results on the selection of the ROI size showed that smaller ROIs can lead to decreased classification due to lack of texture patterns extracted and very large ROIs can include irrelevant features, again resulting in poor classification. The ROI size indicated was comparable with the results shown by Sharma (2017).

## 3.7 Discussion

Due to the variations in datasets, classifiers, the number of images used by researchers, and the differences in density classes chosen for classification, a direct comparison among all the existing methodologies was impossible. Therefore, standard texture feature extraction methods like GLCM were used on the dataset for a comparative study. Similarly, standard LBP variants were compared with the proposed method. Our work mainly used the MIAS dataset but in order to verify the results like ROI size study, we used the DDSM dataset. It is to be noted that the classification results obtained by DDSM was poor which could be attributed to the properties of selected images or due to more variable ground truth annotations as mentioned in (Hadley et al., 2008; Oliver et al., 2008). Similar results can be found in existing literature when using BI-RADS 4 class on DDSM, obtaining a CA of 40.3% using SGLD texture features (Bovis and Singh, 2002), 54.6% using density features (Hadley et al., 2008) and 40.3% using kNN classifier on gray level information (Oliver et al., 2008). So, when correlating our results with the previous studies, it is to be noted that using texture features or gray level information alone is not a best method to classify DDSM into BI-RADS classes. Similarly, it was found that the features obtained by M-ELBP was classifying highly dense and fatty tissue compared to BI-RADS II and III. This could be due to similar glandular structure pattern of II and III classes compared to I and IV. Therefore, additional features (say, morphological features, wavelet features) along with texture could help in better classification. In addition, our study focused more on LBP variants and feature extraction but other directional descriptors like Histogram of Oriented Gradients (HoG) could be investigated in future implementation which may help in extracting the complex glandular tissue pattern features. Similarly, while extracting LBP features for different variants, multi neighbourhood features or boosting features were not extracted as we set our neighbourhood for all LBP variants to be $P = 8$. So,

exploring the multi neighbourhood features may help in extracting more dominant features and can be considered as future work.

## 3.8 Summary

We have developed a new variant of LBP called Mean Elliptical Local Binary Pattern (M-ELBP) by incorporating mean intensity of the neighbourhood along with the binary pattern texture features. Since the method does not perform a direct comparison between the central pixel and neighbouring pixels, it is more robust to noise and illumination changes, unlike the traditional LBP. In addition to introducing a novel variant of LBP, we have investigated its efficiency in breast density classification and found that the LBP variants have the potential of extracting features for density estimation similar to the existing methods. Simultaneously, the effects of selecting ROIs from fibroglandular disc region and selecting the whole breast for classification were studied and was found that selecting ROIs from fibroglandular region can represent the density class of the mammogram image. Similarly, the ROI size and location for classification were studied and found that selecting ROIs of size ranging to $256 \times 256$ pixels from the fibroglandular region was best for representing the density class in our study on MIAS and DDSM data. The results are found to be similar to the existing literature. The classification results by ROI size showed that smaller ROIs can lead to lack of texture patterns to be extracted for classification and larger ROIs can potentially include more irrelevant features decreasing the classification result. Finally, the study on descriptor size showed that there is not much effect in descriptor size while maintaining a balance with the size of the ROI.

# Chapter 4

# Breast Parenchymal Structure Analysis using Blob Modelling

In computer vision, the regions or structures that are identified as either darker or brighter than the local area or surrounding neighbourhood are referred to as blobs (Lindeberg, 1998). Blob detection is a method of visual feature detection of structures or regions. The blob definition varies depending on the purpose as it covers a variety of applications like object detection in navigation (Kiran et al., 2013) and document images (Jetensky, 2015; Swati, 2014), detection of moving objects and vehicles (Bedruz et al., 2017; Jia et al., 2008), traffic sign detection (Zavadil et al., 2012), nuclei detection, and many more. The most common method of estimating the location of blob-like structure is searching for local/global extrema. The size of the structure/blob depends on the scale level at which strongest extremum was detected (Lindeberg, 1998).

In this chapter, a brief description of scale-space theory and automatic scale selection, various blob detection algorithms used in the literature and representation of mammographic parenchymal structures using blob-like structures are presented. In addition, the possibility of detecting blob-like structures from mammographic images to estimate the dense and fatty tissue for risk assessment is explored. Likewise, a novel method of representing mammographic parenchymal patterns through elliptical blob-like structures using scale-space theory is introduced. The possibility of representing parenchymal patterns in elliptical blob forms at different orientation from a real mammogram is explored.

## 4.1 Scale-Space Theory and Automatic Scale Selection

The scale-space theory is a framework of representing and analyzing image structures in multiple stages (Lindeberg, 2013). The scale-space methodology was introduced by (Witkin, 1987) and (Koenderink, 1984) for wide applications in the field of computer vision. Due to the multi-level representation patterns, it is also called as a deep structure. It is a model of perception from the human vision of objects at multiple scales of real-world data. Similarly, it is advantageous in removing unnecessary and confusing details as a method of noise suppression (Lindeberg, 2013). While interpreting the objects in an image, there is no prior knowledge about the size of the object as the level of perception of visualization and interpretation of size may be different. In such cases, the scale-space theory helps in representing and estimating the size of objects in the image by defining the image in different scales. The multi-scale representation of a measured signal is represented by embedding the signal/image into a one-parameter family of derived signal and helps in suppressing the effect of fine-scale perturbations like noise (Koenderink, 1984; Lindeberg, 1993, 2013; Witkin, 1987).

While generating the multi-scale representation of an image, the vital prerequisite is that no new structure should be created during the image transformation process from coarse-scale to fine-scale. Likewise, the object representations at coarse scales should constitute simplifications of respective objects at finer scales (Lindeberg, 2013). The most commonly used convolution kernel for resolution lowering or smoothing is Gaussian (Babaud et al., 1986; Koenderink, 1984; Lindeberg, 1993; Wu and Xie, 1990). When convoluted with a Gaussian kernel, the actual number of pixels in the image does not vary, so the resolution in terms of the number of pixels does not change. But while considering the resolution based on smallest object recognizable details, the convolution process makes the smallest recognizable details larger in image-making a sense that resolution has lowered. Though Gaussian kernel is not the only convolution kernel that lowers resolution without generating additional information, it is the only kernel where we can combine two resolution lowering steps into a single step. This property makes Gaussian kernel the most commonly used convolution kernel in scale space, generating a Gaussian scale space.

The linear scale-space representation of a continuous signal is constructed as follows:- Let $f : \mathbb{R}^N \to \mathbb{R}$ represent any given signal, its scale-space representation $L : \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$ is defined by (Koenderink, 1984; Lindeberg, 1993; Witkin, 1987): $L(\cdot,\cdot;0) = f$ and

$$L(\cdot,\cdot;t) = g(\cdot,\cdot;t) * f(\cdot,\cdot) \tag{4.1}$$

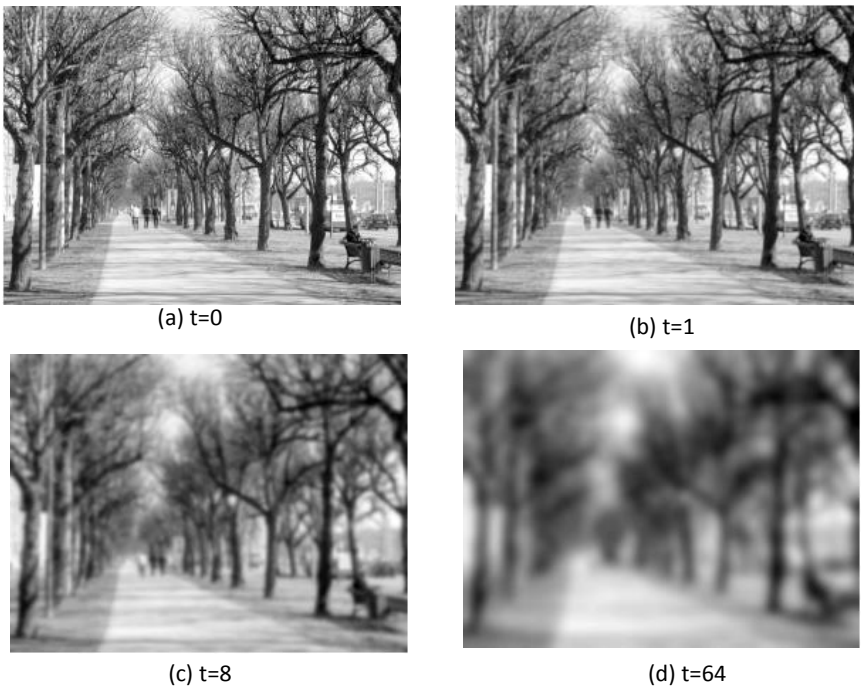



Fig. 4.1 Scale-space representation of an image of size $560 \times 420$ pixels at various scales (figure taken from Lindeberg (2007)).

where $t \in \mathbb{R}_+$ is the scale parameter, and $g : \mathbb{R}^N \times \mathbb{R}_+ \setminus \{0\} \to \mathbb{R}$ is the Gaussian kernel; which could be described in arbitrary dimensions as:

$$g(x;t) = \frac{1}{(2\pi t)^{N/2}} e^{-x^T x/(2t)} = \frac{1}{(2\pi t)^{N/2}} e^{-\sum_{i=1}^{N} x_i^2/(2t)} \quad \left(x \in \mathbb{R}^N, x_i \in \mathbb{R}\right) \tag{4.2}$$

The square root of the scale parameter, $\sigma = \sqrt{t}$, is the standard deviation of the kernel g, which is a measure of spatial scale in the smoothed signal at scale $t$.

While $t = 0$, $L(x,y;0) = f(x,y)$. Thus the scale-space representation of an image at $t = 0$ will be the original image. With the increase in variance $t$, the Gaussian kernel enlarges and removes noisy details from the scale space $L$. For a more detailed example of scale space representation see Figure 4.14.

## 4.2 Blob Detection Algorithms

The most common blob detection algorithms used for various applications in computer vision are template matching, thresholding methods, extrema algorithms, differential algorithms, machine learning and super-pixel classification.

### 4.2.1 Template Matching Algorithms

It is considered as one of the fast and robust methods of detecting blobs/regions in an image. In this method, a template image of the object to be found is overlaid over the search image. On each move, the sum of squared differences (SSD) or sum of absolute differences (SAD) is stored in a correlation matrix. The procedure is continued for the whole search image. The highest values in the correlation matrix represent the local maxima which indicate the highest probability of object location in the search image. Later the exact number and location of blobs can be extracted using a threshold value. Though it is considered as a fast method, it is difficult to find objects that do not match with the precise template (Kaspers (2011), Sezgin and Sankur (2004), Boyle et al. (2008)).

### 4.2.2 Thresholding Algorithms

While considering blobs as bright or dark regions in an image, thresholding helps in attaining binary images. The holes in the blob regions are filled up initially, and then the group of neighbouring pixels are searched using connected components to detect the related regions. Later, the location of the blob/object is estimated by computing the centroid of the blob. Sezgin and Sankur (2004) provided a detailed survey of different thresholding methods and their applications.

### 4.2.3 Local Extrema

It is the procedure in which the local minimum or local maxima of the bright regions or objects in an image is estimated. During run-time for a window size (say $3 \times 3$), the local maxima and minima above a threshold value are recorded as the location of the centre of blob. Later, a filtering process over the detected points helps in removing the extrema that are not the centre of the blobs. The major drawback of this technique is that it fails when there are multiple extrema in a single region.

### 4.2.4 Differential Extrema Calculation

The most commonly used method in blob detection using scale-space representation. It finds a wide range of application in medical image processing and related applications. Differential extrema calculation find its application better when local extrema algorithms fail to find local extrema due to noise. The differential extrema methods are based on the derivative of intensity functions. Most of the differential extrema

methods use the scale-space analysis for blob location and size determination. In this section, the most common blob detection methods based on scale-space are explained.

**Laplacian of Gaussian (LoG):**

This is the most popular differential method of blob detection based on the Laplacian of Gaussian (LoG) operator (Lindeberg, 2013). The method first convolves the search image with a Gaussian kernel at a specific scale $t = \sigma^2$ generating the scale-space representation $L(x,y;t)$, where $x$ and $y$ are the corresponding pixel coordinates. Then it applies the Laplacian operator to generate a strong response at the blob regions. The blobs are detected at scale-space extrema of the normalized Laplacian $\nabla^2_{\text{norm}} L$ which is the local extrema with respect to spatial location and scale level (Lindeberg, 1998).

The normalized Laplacian $\nabla^2_{\text{norm}} L$ is defined as :

$$\nabla^2_{norm} L = \sigma^2 (L_{xx} + L_{yy}) \tag{4.3}$$

where $L$ is the scale space representation of image $I(x,y)$ through convolution by Gaussian kernel at various scales $t$:

$$L(x,y;\sigma) = G(x,y;\sigma) * I(x,y) \tag{4.4}$$

where * represent the convolution operation and $G(x,y;\sigma)$ defined as

$$G(x,y;\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{4.5}$$

The Laplacian operator $\nabla^2$ is applied to the Gaussian scale-space representation of an image as

$$\nabla^2 = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{4.6}$$

or the LoG operator 4.7 is computed and then convoluted with the image to generate LoG scale-space representation.

$$\nabla^2 G(x,y) = \frac{x^2+y^2-2\sigma^2}{\pi\sigma^4} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{4.7}$$

The scale levels at which the extrema is strongest will define the size and location of the blob.

Due to the scale-space detection of blobs at different scales, it finds many applications like detecting grey level blob-like structures from images (Lindeberg, 1993), cell

detection from microscope images (Samborskiy, 2007), registering histological and MR images of prostate for cancer detection (Zhan et al., 2007), nuclei detection (Kong et al., 2013a), detection of spherical lesions in 3D medical images (Fazlollahi et al., 2013), detection of lung nodule (Saien et al., 2014), and detection of multiscale blob-like structure from mammograms (Chen et al., 2013b).

**Difference of Gaussian (DoG)**

Similar to LoG, in DoG, the image is first smoothened through convolution by Gaussian kernel of certain size $\sigma_1$.

$$G_{\sigma_1}(x,y) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{x^2+y^2}{2\sigma_1^2}\right) \tag{4.8}$$

to generate

$$g_1(x,y) = G_{\sigma_1}(x,y) * f(x,y) \tag{4.9}$$

again using a different scale $\sigma_2$, a second smoothed image can be obtained through convolution as:

$$g_2(x,y) = G_{\sigma_2}(x,y) * f(x,y) \tag{4.10}$$

The detection of scale-space extrema of an image $f(x,y)$ can be obtained by calculating the extrema from the Difference of Gaussian function $DoG * f(x,y)$ as:

$$g_1(x,y) - g_2(x,y) = G_{\sigma_1} * f(x,y) - G_{\sigma_2} * f(x,y) = (G_{\sigma_1} - G_{\sigma_2}) * f(x,y) = DoG * f(x,y) \tag{4.11}$$

Though the DoG is faster than LOG in computation, both the operators have a drawback that it will detect strong local extrema near the edges in an image where the relative intensity changes in the spatial direction.

**Determinant of the Hessian Matrix (DoH)**

It is considered as a faster method compared to LoG and DoG and is independent of the size of blob size, as it uses box filters instead of convolutions. Compared to Laplacian, it has better scale selection under linear image transformations. It detects the blobs by calculating the maximums in the matrix of DoH.

The Hessian matrix of a scale-space image representation $L$ is defined as

$$H = \begin{bmatrix} L_{xx} & L_{xy} \\ L_{yx} & L_{yy} \end{bmatrix} \tag{4.12}$$

Based on the studies by Lindeberg (1998), the blobs at the scale-space representation can be detected from the determinant of the normalized Hessian matrix of $L$ as:

$$\det H_{\text{norm}} L = \sigma^2 \left( L_{xx} L_{yy} - L_{xy}^2 \right) \tag{4.13}$$

The radius of the blobs is proportional to the scales of image representation. Unlike LoG and DoG, the DoH can eliminate the extrema at the edges and contours where the intensity changes in single spatial direction, as it detects the blobs based on the determinant of the Hessian matrix.

## 4.3 Synthetic Models to Represent Breast Parenchymal Patterns

Parenchymal patterns can be considered as building blocks of mammographic tissue modelling. It plays an important role in mammographic risk assessment and density assessment. As per Tabár's (Tabar et al., 2003) definition of breast parenchymal structure, the breast is made of four building blocks: nodular, linear, homogeneous and radiolucent (see Section 1.4 for more details). In our study, we attempt to model the mammographic tissue using blob model and attempt to classify them as high risk and low-risk mammograms. There are different mammographic synthetic models developed for representing the breast tissue structure. Bochud et al. (1999) tried to model the background tissue to generate realistic simulated mammogram images. Bochud et al. (1999) used the images obtained by images called lumpy backgrounds generated by Rolland and Barrett (1992). The lumpy backgrounds are found to be more statistically tractable and stationary. The lumpy backgrounds consist of a random number of structural elements called blobs, located at random locations. Bochud et al. (1999) used a generalised clustered lumpy background by clustering the blobs within the image. The blob clustering helped in generating textures which are visually similar to the real images with similar histograms and power spectra.

Similarly, Heine et al. (1999) developed a simulated model where many statistical characteristics resembled real mammograms. The model was developed by the process of passing a random input field through a linear filter with self-similar feature. The model gave an explanation of the image appearance and the grey level distribution pattern typical in mammograms. Though the model was a success in describing the multi-model grey value distribution, it failed to relate the 3D tissue structures like the duct patterns.

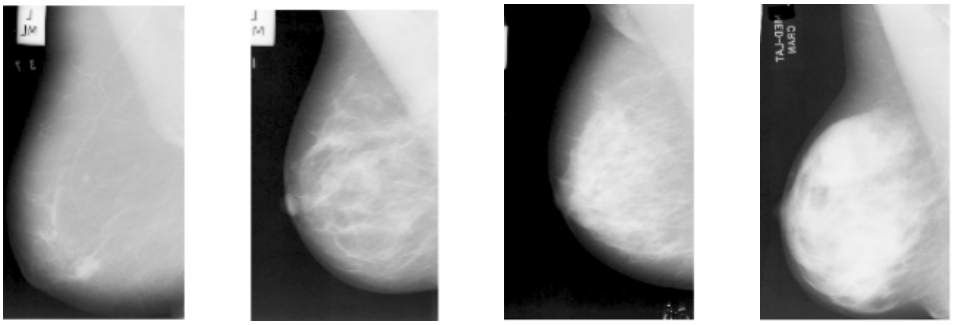
Taylor et al. (1998) developed a simulation model to explain the duct system in the breast, based on the fractal features using the information on duct length and diameter. It helped in providing a simulated representation of the growth of duct systems within the breast volume.

Bakic et al. (2002a) later generated the synthetic mammograms based on 3D simulation of mammography. Simulated models of 3D breast anatomic structures were projected towards to form a synthetic mammogram image. The model successfully generated a realistic distribution of large and medium scale distributions. The mammographic modelling process was simulated using a compression model which estimates the breast deformations using tissue elasticity parameters and an x-ray image acquisition process where a monoenergetic parallel beam was applied to synthetically compressed phantom. Later, Bakic et al. (2002b) analysed the synthetic mammogram texture to real mammogram images. The synthetic images were developed by using x-ray image acquisition method on computationally compressed phantoms. The phantoms represented the different texture features in the images and were generated by simulating different size of structural elements like spherical/ellipsoid shells and blobs. The study showed that the model was successful in developing synthetic images with a similar distribution of features like the average size of the image objects, the texture energy and the fractal dimension with the real mammogram images. Though the model was successful in generating medium-sized phantoms closely related to the mammographic object size, it was less capable in modelling small structures like fine tissue details, lymph vessels etc.

In our work, we attempt to estimate the breast density and risk by generating a blob model based on the parenchymal structures. We utilise the scale-space theory using LoG to generate the scale-space representation and to estimate the scale size of dense regions in the breast. Chen et al. (2013b) detected blob-like tissue patterns from mammographic images to estimate the breast density and found a high correlation with the Breast Imaging Reporting and Data System (BI-RADS) density classification. Using Chen et al. (2013b) blob-like tissue pattern detection model and the Bakic et al. (2002b) synthetic mammographic model generated by projecting various phantoms structures like spherical shells and blobs, we attempt to find the orientation of tissue structures through an elliptical blob model.

## 4.4 Mammographic Blob Modelling Towards Risk Estimation

Breast tissue density and parenchymal pattern structures are considered to be an indication of high risk and they have a direct relation with the probability of developing breast cancer. A few examples of mammographic images are illustrated in Figure 4.2 from MIAS dataset clarifying various types of tissue densities as per the BI-RADS density classification.



(a) (b) (c) (d)

Fig. 4.2 Some example mammograms showing various BI-RADS density classification based on BI-RADS 4th edition. (a) BI-RADS 1 (entirely fatty breast tissue), (b) BI-RADS 2 (scattered fibro-glandular density), (c) BI-RADS 3 (heterogeneously dense breast obscuring small masses), and (d) BI-RADS 4 (extremely dense breast lowering the sensitivity of mammography).

As an initial step, we attempt to model the fatty and dense tissue in the mammograms to investigate the possibility of finding the relationship between modelling parenchymal patterns in relation with BI-RADS class. This work is closely related to that of Chen et al. (2013a), which developed a mammographic segmentation approach using topographic maps with an emphasis on modelling and distribution of bright blobs in mammographic images using LoG operator. As traditional methods emphasis on detection of circular blob-like structures from the patterns, approaches like LoG and DoG filtering have been used. But images can have asymmetrical patterns like in mammograms. So, to facilitate detection of orientated ellipse-like structures in images, Kong et al. (2013a,b) introduced a generalized Laplacian of Gaussian (gLoG) based approach, which allows the detection of symmetric and asymmetric regions and includes direction estimation for such regions. The gLoG approach uses a set of orientation and scale-dependent kernels. As with the LoG approach, the image, $f(x,y)$

is processed with a Gaussian kernel and differential aspects as defined by

$$\nabla^2[G(x,y) * f(x,y)] = \nabla^2[G(x,y)] * f(x,y) \tag{4.14}$$

where

$$\nabla^2 G(x,y) = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2} \tag{4.15}$$

and

$$G(x,y) = A \cdot e^{-\left(ax^2+2bxy+cy^2\right)} \tag{4.16}$$

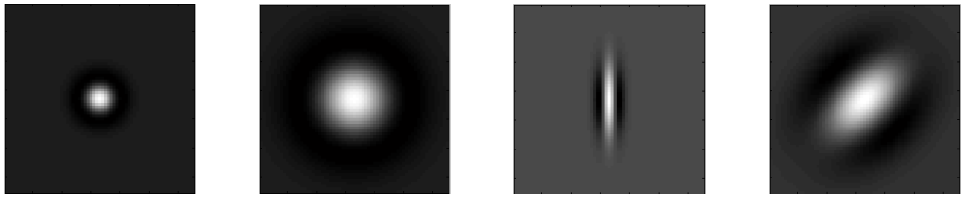
where the parameters $a, b, c$ are defined as

$$a = \frac{\cos^2\theta}{2\sigma_x^2} + \frac{\sin^2\theta}{2\sigma_y^2} \tag{4.17}$$

$$b = -\frac{\sin 2\theta}{4\sigma_x^2} + \frac{\sin 2\theta}{4\sigma_y^2} \tag{4.18}$$

$$c = \frac{\sin^2\theta}{2\sigma_x^2} + \frac{\cos^2\theta}{2\sigma_y^2} \tag{4.19}$$

where the $\sigma_x$, $\sigma_y$ and $\theta$ are respectively the standard deviations in the horizontal and vertical directions that determine the scale/width of the Gaussian kernel, and the orientation of Gaussian kernel. If $\sigma_x$ and $\sigma_y$ carry the same standard deviation, we get a symmetrical shaped LoG operator. Some typical kernels can be found in Figure 4.3.



(a) (b) (c) (d)

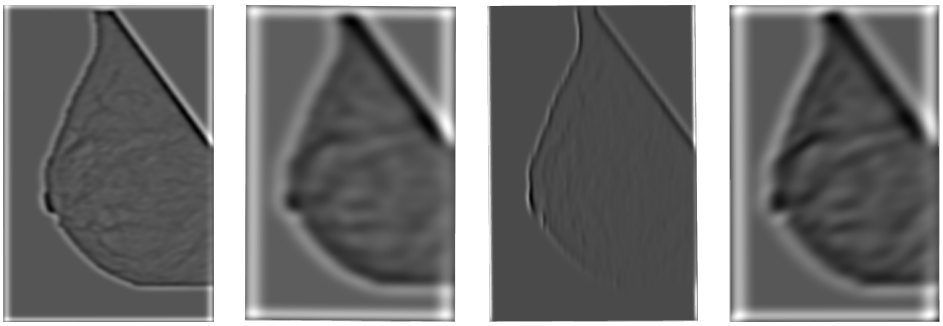
Fig. 4.3 A set of gLoG example kernels, where (a) $\{\sigma_x = 4, \sigma_y = 4\}$, (b) $\{\sigma_x = 10, \sigma_y = 10\}$, (c) $\{\sigma_x = 8, \sigma_y = 2, \theta = \pi/2\}$, and (d) $\{\sigma_x = 12, \sigma_y = 8, \theta = \pi/4\}$.

### 4.4.1 Data and Experimental Set-up

For the initial study, we use mammogram images from the MIAS dataset (refer to Section 2.9.1 for more details on the dataset). The BI-RADS density type of all the 320 images from the dataset was marked by expert radiologists (2 images from the

database were excluded due to corruption). All the images are pre-processed, and the breast area is separated from the background region and the pectoral muscle.

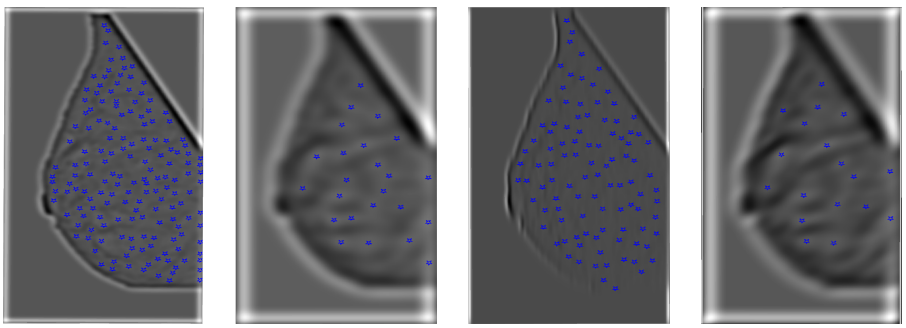
The blob-like features from the images at scale-space using gLoG operator are extracted. A few typical examples can be found in Figure 4.4 which shows that different size and orientation structures are localized and enhanced by other kernels. Blob-like structures detected outside the breast area are removed by the erosion of the breast mask and combining it with the response image.



(a) (b) (c) (d)

Fig. 4.4 Example gLoG processing based on the mammogram shown in Figure 4.2 (b) and the gLoG kernels shown in Figure 4.3.

Based on the gLoG results we identified local maxima, which indicates the central location of blob-like structures and thresholding could be used to select only those regions with a higher probability. Some blob detection results can be found in Figure 4.5.

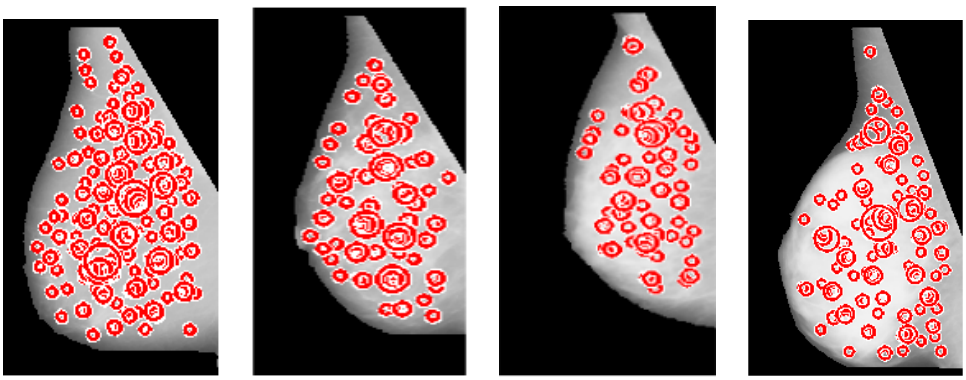


(a) (b) (c) (d)

Fig. 4.5 Detected blobs super-imposed on the original mammogram based on the images shown in Figure 4.4.

In order to study the parenchymal arrangement of breast tissue through blob modelling, we used the scale-space theory with gLoG operator for $\sigma_x = 10$, $\sigma_y = 10$

and $\theta = 0$ to exclude the orientation effect as an initial study. The extrema of local regions in the image at different scales are estimated based on the highest probability regions in each scale. Such detected blobs represent the approximate size of blob-like fatty and dense tissue regions. To estimate the amount of blob-like fatty/dense tissue structures in the image, the blobs are identified at multiple scales. All these blobs are merged to represent the total amount of dense/fatty tissue in the image. On merging, blob overlapping occurs in areas where tissues are closely related. Blobs detected for fatty tissue from all the scales are represented in Figure 4.6 and the problem of overlapping blobs is evident. This is because the same blob-like tissue structures are detected as blobs at different scales with their origin at a similar location as they have high probability regions in every scale. This problem of overlapping is solved by distance/radius based blob merging based on the qualitative relations between the blobs. The qualitative relations between two blobs (blob A and blob B) under consideration are the distance ($d$) between the centre points of blobs and the radius of each blob ($r_A$ and $r_B$, where $r_A$>$r_B$). The three common relations found while merging are: external where the distance between the blobs is greater than the sum of its radius($d$>$r_A$+$r_B$), internal ($d$<$r_A$-$r_B$) and intersection ($r_A$-$r_B$ <$d$<$r_A$+$r_B$).

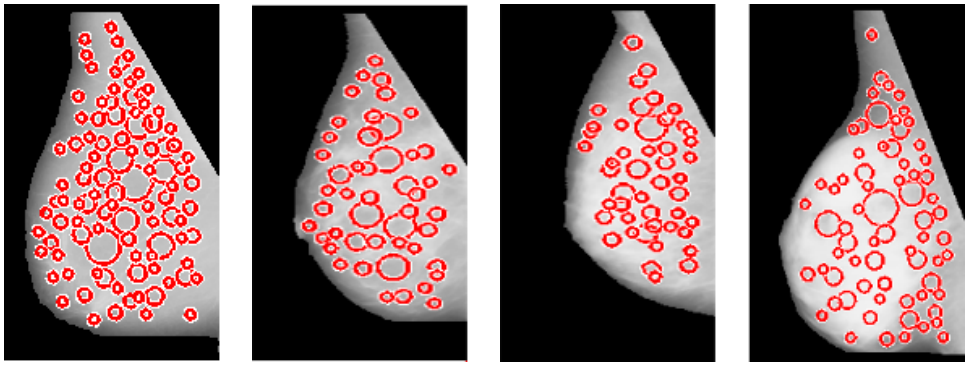


(a) (b) (c) (d)

Fig. 4.6 Blobs detected fatty tissue at multi-scale for mammograms sorted from BI-RADS I to BI-RADS IV as per Figure 4.2.

The overlapped and intersected blobs are removed using the above qualitative relations. The merging procedure starts from the largest scale to the smallest scale. The external blobs are retained. If the distance between the blobs is less than the radius of the largest blob, the inner blob is eliminated. If they are intersecting and the distance between them is less than the radius of the largest blob, it leads to its deletion. The final merged multi-scale blobs after the removal of additional blobs are illustrated

in Figure 4.7. The corresponding tissue area is normalized using the total breast area for the final estimation of the tissue area.

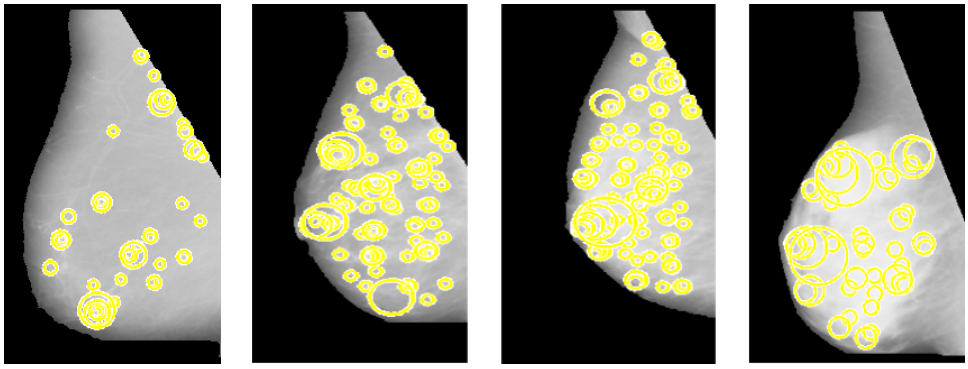


(a) (b) (c) (d)

Fig. 4.7 Blobs detecting fatty tissue at multi-scale after merging based on qualitative relations for mammograms sorted from BI-RADS I to BI-RADS IV.

.

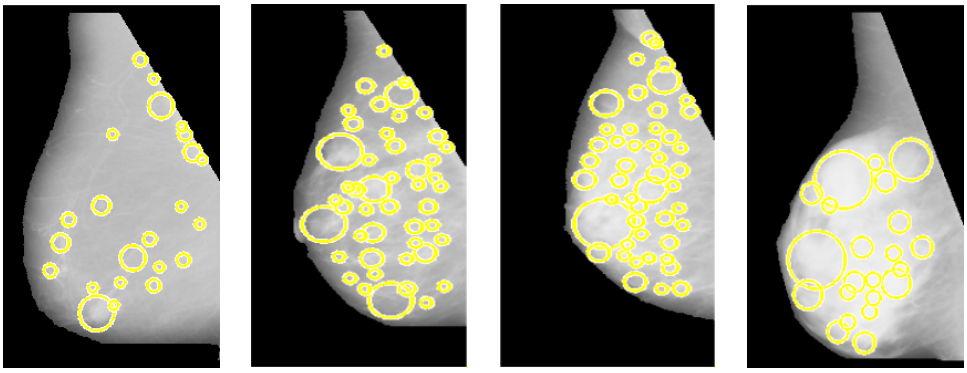
As discussed before the same procedure can be used for estimating the blob-like dense tissue structures in a mammogram image by inversing the kernel and identifying local extrema. The sample images are shown in Figures 4.8 and 4.9. Figure 4.8 illustrates the blobs detected at all scales and Figure 4.9 illustrates the dense blob-like structures where the overlapped and intersected blobs have been removed through the qualitative measurements.



(a) (b) (c) (d)

Fig. 4.8 Blobs detected dense tissue at multi-scale for mammograms sorted from BI-RADS I to BI-RADS IV.

It is clear from the Figures 4.7 and 4.9 that the number of blobs detecting fatty tissue area decreases, while the amount of dense tissue area increases for BI-RADS I to BI-RADS IV due to the increase in glandular tissue. The result matches with



(a) (b) (c) (d)

Fig. 4.9 Blobs detecting dense tissue at multi-scale after merging based on qualitative relations for mammograms sorted from BI-RADS I to BI-RADS IV.

BI-RADS description that in BI-RADS I category, the breast tissue is entirely fatty while the glandular tissue pattern increases with BI-RADS category level. Therefore, the amount of dense blobs is more in BI-RADS IV relative to fatty blobs. While considering fatty tissue, the number of blob-like fatty tissue covering breast region will be more in BI-RADS I category as most of the breast region are occupied by fatty tissue than dense. Likewise, the number and size of blobs representing dense tissue will be more towards BI-RADS Iv category as majority of breast region is covered with dense tissue compared to fatty tissue.

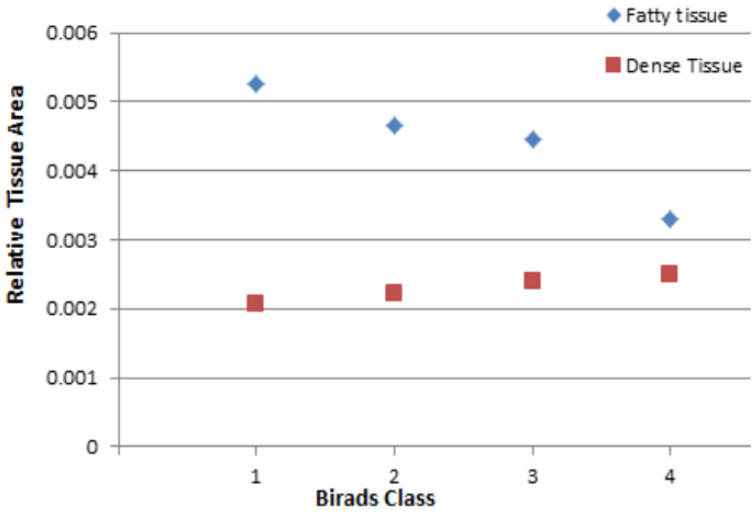



Fig. 4.10 Average relative tissue area for all the MIAS mammograms.

While comparing the relationship between dense and fatty tissue in estimating BI-RADS class, it is found from Figure 4.10 that when relative dense tissue increases with BI-RADS class, the fatty tissue area decreases with BI-RADS class in correlation

with the dense tissue. Besides, it shows that there is a close relationship between fatty tissue area for BI-RADS II and BI-RADS III indicating that the fatty tissue pattern for BI-RADS II and BI-RADS III is nearly similar compared to the drastic change in structure between BI-RADS I and BI-RADS IV.

Although the Figure 4.10 shows an average linear increase of dense tissue and decrease of fatty tissue respectively with BI-RADS class, there is significant variation in the dense/fatty tissue. This could be due to the over/under segmentation of tissue areas. Further investigation is needed to control the over/under segmentation while detecting the fatty/dense tissue area through multi-scale blobs.

Similarly, to study the potential of blob modelling in risk assessment, we classify the dense and fatty blob regions using the classical KNN classifier. The result shows an overall classification of 44.69% with specificity, sensitivity and F-score of 0.46, 0.45 and 0.45, respectively. The confusion matrix for the KNN classifier is represented by Table 4.1.

Table 4.1 Confusion matrix for KNN classifier based on BI-RADS classes.

| | | **Automatic Classification** | | | | |
|---|---|---|---|---|---|---|
| | | **BI-RADS I** | **BI-RADS II** | **BI-RADS III** | **BI-RADS IV** | **CA%** |
| **Truth Data** | **BI-RADS I** | 44 | 26 | 17 | 0 | 50.6 |
| | **BI-RADS II** | 20 | 55 | 24 | 4 | 53.4 |
| | **BI-RADS III** | 11 | 44 | 33 | 5 | 35.5 |
| | **BI-RADS IV** | 1 | 14 | 11 | 11 | 29.7 |

To estimate the risk in density, we categorize the mammogram images in MIAS dataset into two density groups: high density and low-density group. The images are also classified into high/low risk using classical KNN classifier using a ten-fold cross-validation. The classification results, as shown in Table 4.2 indicate a correlation to the Table 4.1 with a classification accuracy of 66.31%, specificity of 0.66, sensitivity of 0.66 and F-measure of 0.65.

Table 4.2 Confusion matrix of KNN classifier for classifying low/high risk.

| | | **Automatic Classification** | | |
|---|---|---|---|---|
| | | **Low Risk** | **High Risk** | **CA%** |
| **Truth Data** | **Low Risk** | 156 | 34 | 82.1 |
| | **High Risk** | 74 | 56 | 43.1 |

### 4.4.2 Discussion

The initial study based on circular blobs for classification of mammograms into BI-RADS classes through blob modelling shows the potential of ellipse modelling. The

Figure 4.10 shows the distribution of fatty/dense tissue in four categories but the classes II and III seem to be similar which could be due to the related structure of parenchymal distribution as the glandular tissue are heterogeneously spread in the whole breast region. Therefore, a more detailed scale-space structuring is necessary to distinguish the minor differences in the tissue arrangement.

Similarly, the study on risk assessment showed the possibility of further investigation of blob modelling in risk assessment with a classification accuracy of 66.3%. The initial research leads the pathway for more complex elliptical blob modelling for mammographic risk assessment which will be explained in detail in the following sections.

## 4.5 Representation of Mammographic Parenchymal Structures using Elliptical Blob Model

As discussed in the previous section and studies on synthetic mammogram generation, the parenchymal pattern modelling using geometrical structure like blobs was found to be useful in risk assessment and density estimation. So, to improve the modelling, we attempt to explore an elliptical model for representing and calculating the parenchymal patterns. The work was based on the initial study by Chen et al. (2013b) where the modified LoG filter was used to reduce the boundary effect caused due to Laplacian convolution. Figure 4.11 shows the possibility of representing the parenchymal patterns in mammograms using different scales of a blob.

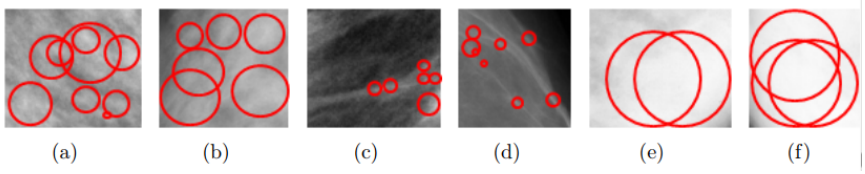



Fig. 4.11 Pseudo blobs superimposed on image patches of example breast tissue patterns: (a), (b) nodular tissue; (c), (d) linear tissue; (e), (f) homogeneous tissue. (Image taken from Chen et al. (2013b).

.

Though there are numerous methods in the literature for representing mammographic parenchymal patterns, none of the methods has modelled them in multi-scale elliptical blob model in various directions. Our method focuses on the spatial distribution of elliptical blob-like salient regions in mammographic breast region in multi-scale structure. We perform a gLoG approach to extract even the smallest blob-like regions.

### Pre-processing

We estimate the parenchymal breast patterns in the breast region; therefore, the pectoral muscle and artefacts like labels are to be removed from the mammogram. The pectoral muscle and label artefacts are to be removed as the initial procedure, as it could produce strong responses during the convolution and extrema calculation from scale-space. Likewise, the pectoral muscle intensity will be higher compared to parenchymal breast patterns which can cause deletion of blob regions while estimating the threshold. So, to prevent such cases, we remove the artefacts and pectoral muscle regions and focus on the patterns inside the breast region. To remove pectoral muscle, we use breast segmentation method proposed by Chen and Zwiggelaar (2010). But breast region segmentation can cause sharp edges through breast boundaries like regions near to pectoral muscle (G, F), regions to the edges of mammogram image border (I, H, A, B, C, D, E) and breast borders as shown in Figure 4.12.

To circumvent this, we use a generalized model of LoG; it will generate responses not only to the salient regions but also to edges and borders as they bring edge effect during the convolution process in scale space generation. To reduce the edge and border effect, the images were filtered using a modified LoG operator proposed by Chen et al. (2013b) where the filter template alter adaptively according to the local window to be convoluted with it. Therefore, the defined mask ($M(x,y)$) restrict the filtering to non-zero pixels within the local window under convolution.

If pixel $I(x,y) \neq 0, M(x,y) = 1$; else $M(x,y) = 0$. The normalization of the filtering operators helps in reducing the response to homogeneous regions in an image. Therefore, the modified LoG operator is normalized based on the mask area as

$$f_m(x,y) = \begin{cases} 0 & \text{if } M(x,y) = 0 \\ f(x,y) - \sum_{(x,y)\in T \& M(x,y)-1} f(x,y)/N_M & \text{if } M(x,y) = 1 \end{cases} \tag{4.20}$$

where, $N_M$, the number of non-zero pixels within the local window under convolution, computed by $N_M = \sum_{(x,y)\in T} M(x,y)$ and $f(x,y)$ is the LoG operator. Thus, the filter coefficients within the mask area will be normalized to zero.

In addition, we create the elliptical gLoG kernels for different orientations $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$. The same approach as discussed in Section 4.4 is used to develop the kernels. A sample of the elliptical kernels are represented in Figure 4.13.

The kernels at different orientations will be used later for estimating the extrema produced by parenchymal patterns as per their orientation.

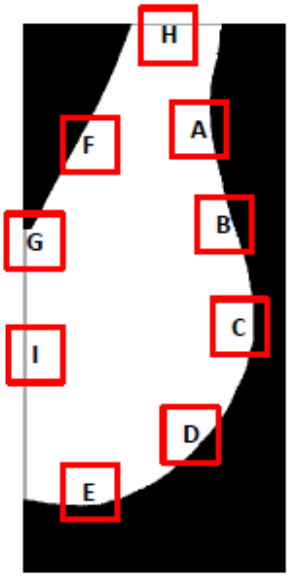



Fig. 4.12 Possible sharp edges and borders after pre-processing. (image from Chen et al. (2013b)).

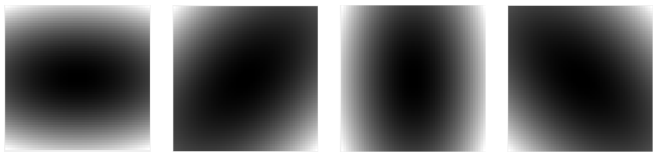


Fig. 4.13 Examples of elliptical Laplacian of Gaussian kernels at different orientation (orientations from $\theta = \{0°, 45°, 90°, 135°\}$ from left to right).

### 4.5.1 Ellipse Blob Detection

After pre-processing, the scale-space representation of mammogram images is generated for the local extrema and global extrema calculation. Instead of enlarging the operator size for each scale, we scale to reduce the mammogram image. Otherwise, repeated convolution with large kernels with images at original scales will be computationally complex. In order to subsample the image for scale-space generation, we use the sub-sampling factor of $1/\sqrt{2}$ for a scale level of 10, starting with the first scale with $\sigma_{x1}$=8 and $\sigma_{y1}$=$\sigma_{x1} * 1.5$. A sample of generated scale space is represented in Figure 4.14.

The initial image down-sampled at a factor of 1/k is then filtered using gLoG operator at various orientation. The filtered images are then up-sampled to the original size of the image and squared for calculating the scale-space extremum, as shown in Figure 4.15.

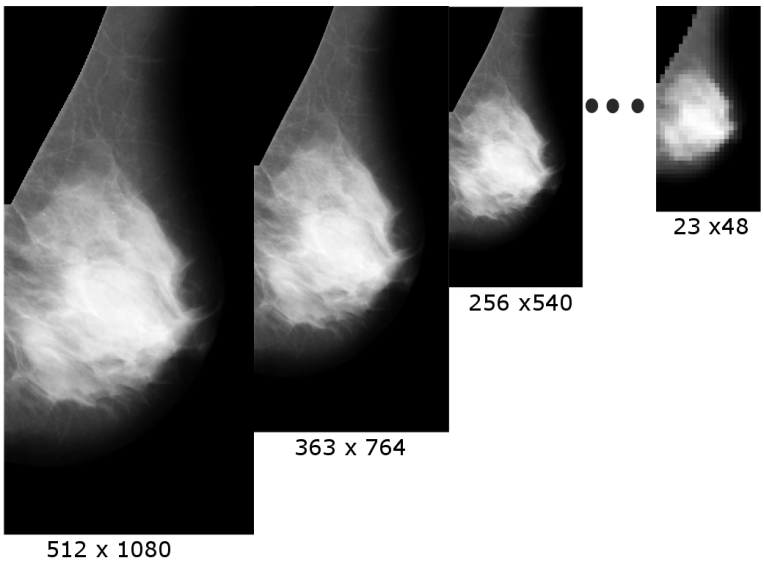



Fig. 4.14 Scale space representation of mammographic images at different scales.

Unlike Section 4.4.1, where both fatty and bright blobs were considered, here we focus on the bright tissue regions from the mammographic images. The bright regions in mammograms represent the dense tissue, and they are described as bright blobs in scale-space (represented as negative filter response).

To search the extrema in scale-space, we first detect the extrema in each scale and then compute the global extremum. For computing the local extremum, each image pixel is compared with its local neighbourhood with a window size of $5 \times 5$ region. The local extremum response with low probability/contrast is discarded based on a threshold value. The threshold value is computed based on the maximum response value of the pixel in the filtered image. After removing the less probable responses, the extrema are stored as they indicate the centre of the blob at each scale where the scale represents the minor axis radius for the blob at a respective orientation. Since we concentrate on the patterns inside the breast region, the elliptical blobs which go outside the breast boundary are rejected based on area. If the majority of blob region is outside the breast region, it will be rejected from the blob pool. A sample of local extrema detection at different scales is shown in Figure 4.16.

The blob structures at each scale after removing the blobs having mean intensity less than the threshold is selected as salient blob regions at the respective scale. A sample of selected blobs at each scale and orientation is represented in Figure 4.17.

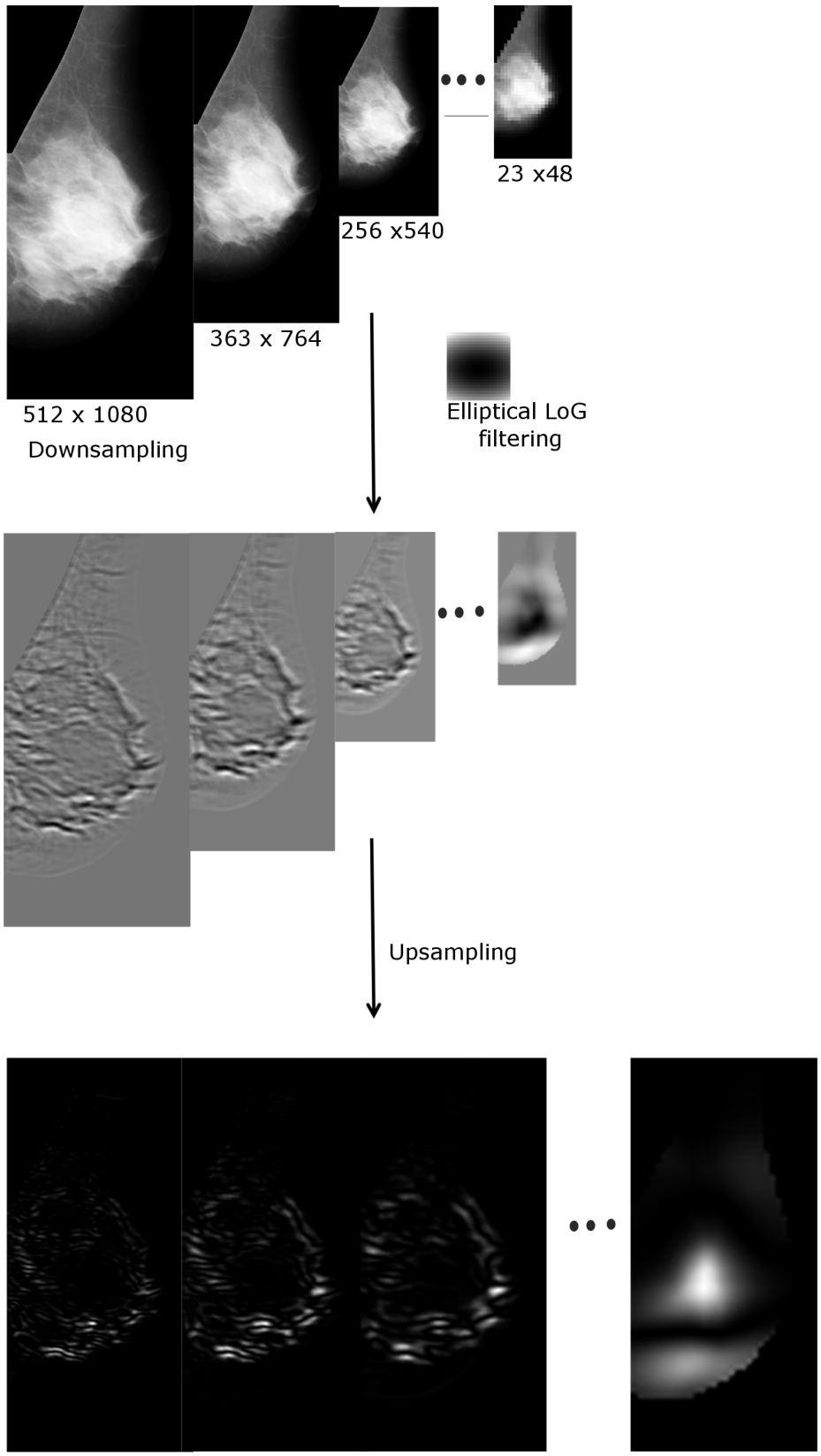



Fig. 4.15 Level 1: Scale space representation of mammographic images at different scales as in Fig 4.14, level 2: elliptical LoG filtered images at 0°, level 3: squared elliptical LoG filtered responses on upsampled images.

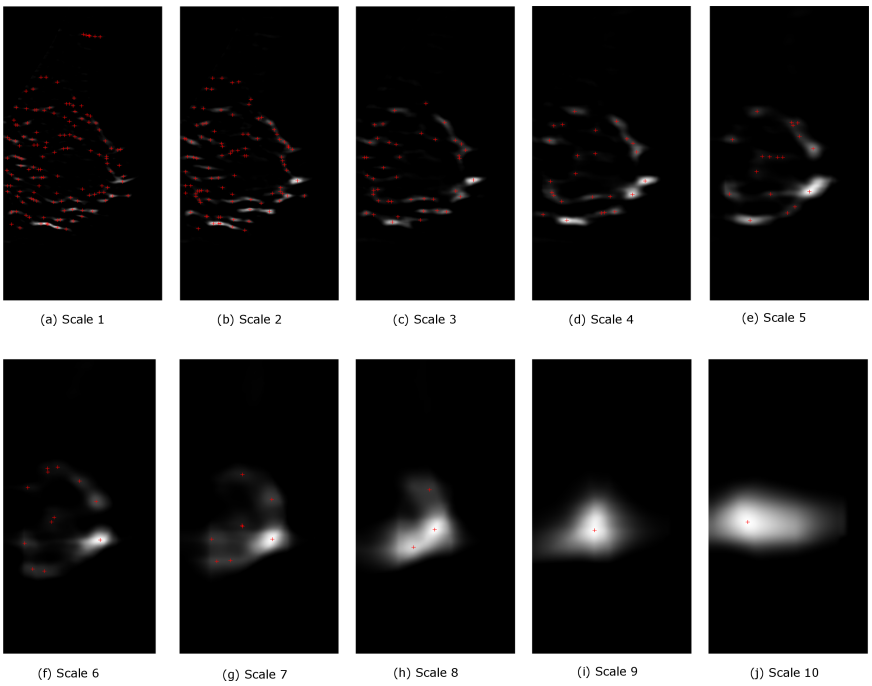


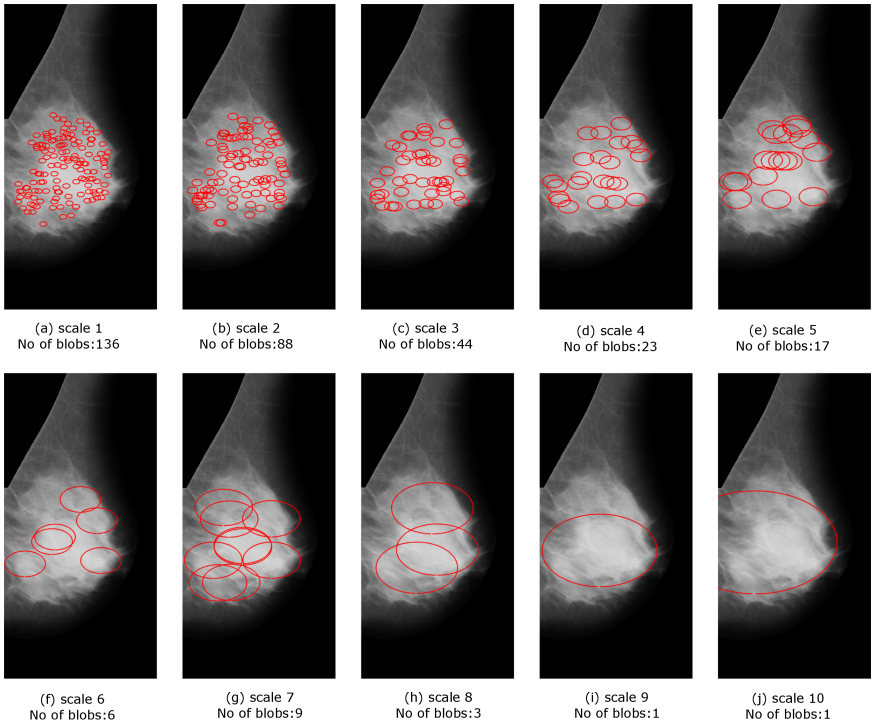

Fig. 4.16 The squared gLoG filter response for a mammogram image (shown in Fig 4.15) at ten different scales for an orientation of 0°. The positive responses and less probable extrema have been removed.

After detecting the local extremum at different scales, the most robust responses across the scales are estimated for each orientation. Blobs at each scale are compared with its corresponding extremum in other scales by comparing with its $5 \times 5 \times 9$ neighbours (from different scales). Finally, only the global extremum among all the scales are retained for each orientation. The resulting blobs after global extremum are shown in Figure 4.18 top row.

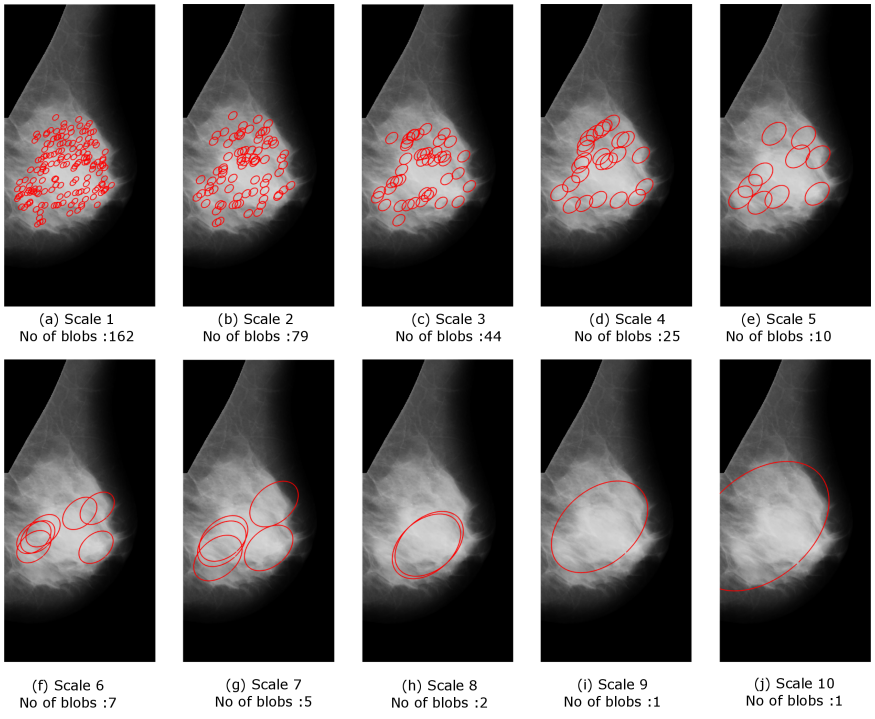
### 4.5.2 Blob Merging

The blobs detected at different scales represent the dense tissue patterns inside the breast region at various orientations. The blobs representing dense tissue may overlap at different scales due to its intensity or by the size of the region it covers. Therefore, to estimate the dense tissue covered by blob regions, we merge the overlapping blobs as in Figure 4.18 (bottom row).

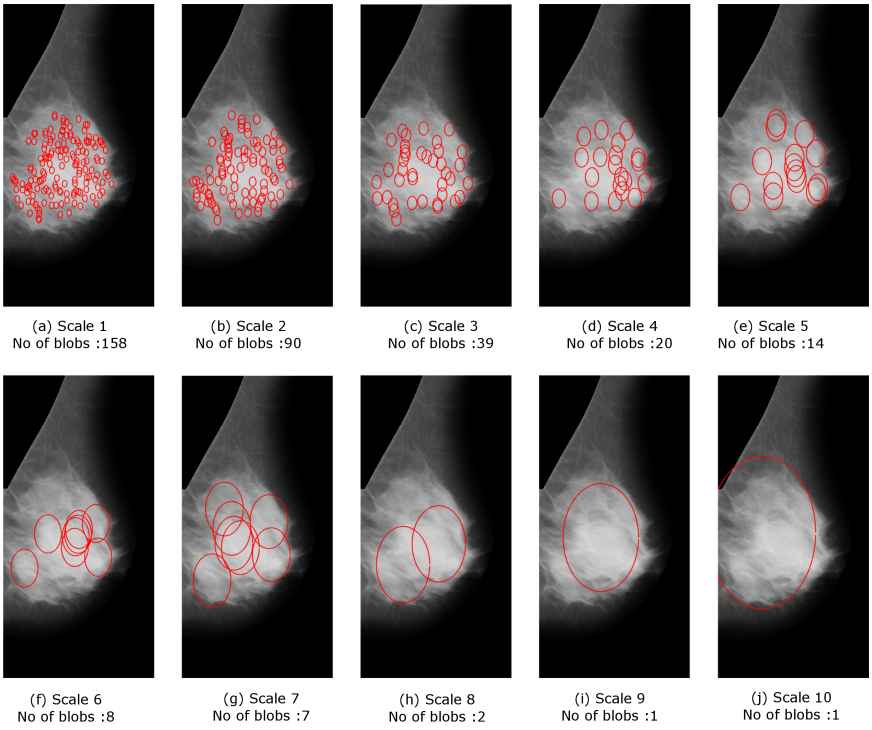
For blob merging, a direct distance-based approach based on the radius is not applicable due to the elliptical structure of the blob and the multiple orientations. Therefore, we use an area-based merging method for deleting the overlapping blobs. Suppose, there are two elliptical blobs A and B, where A is larger than B. If A completely overlaps B, then B is deleted as shown in Figure 4.19 (a). If A and B


(a) scale 1
No of blobs:136
(b) scale 2
No of blobs:88
(c) scale 3
No of blobs:44
(d) scale 4
No of blobs:23
(e) scale 5
No of blobs:17
(f) scale 6
No of blobs:6
(g) scale 7
No of blobs:9
(h) scale 8
No of blobs:3
(i) scale 9
No of blobs:1
(j) scale 10
No of blobs:1

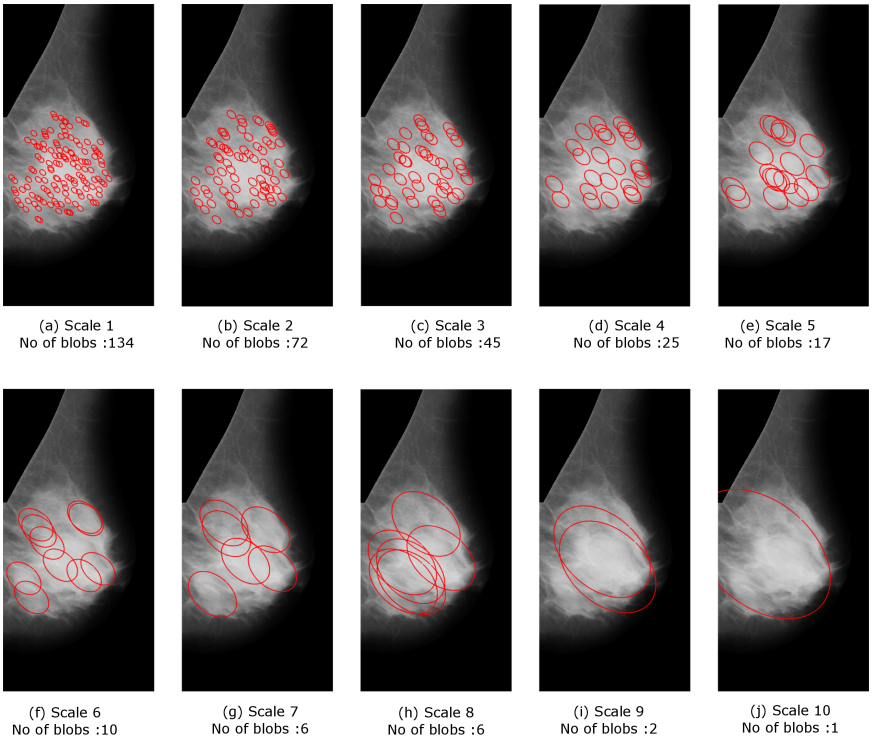
(a) i


(a) Scale 1
No of blobs :162
(b) Scale 2
No of blobs :79
(c) Scale 3
No of blobs :44
(d) Scale 4
No of blobs :25
(e) Scale 5
No of blobs :10
(f) Scale 6
No of blobs :7
(g) Scale 7
No of blobs :5
(h) Scale 8
No of blobs :2
(i) Scale 9
No of blobs :1
(j) Scale 10
No of blobs :1

(b) ii




(a) iii




(b) iv

Fig. 4.17 Detected blobs at each scale after removing the false positives with mean intensity value less than the threshold. ((a)i) blobs at different scales for 0° orientation; ((b)ii) blobs at different scales for 45° orientation; ((a)iii) blobs at different scales for 45° orientation; ((b)iv) blobs at different scales for 135° orientation.

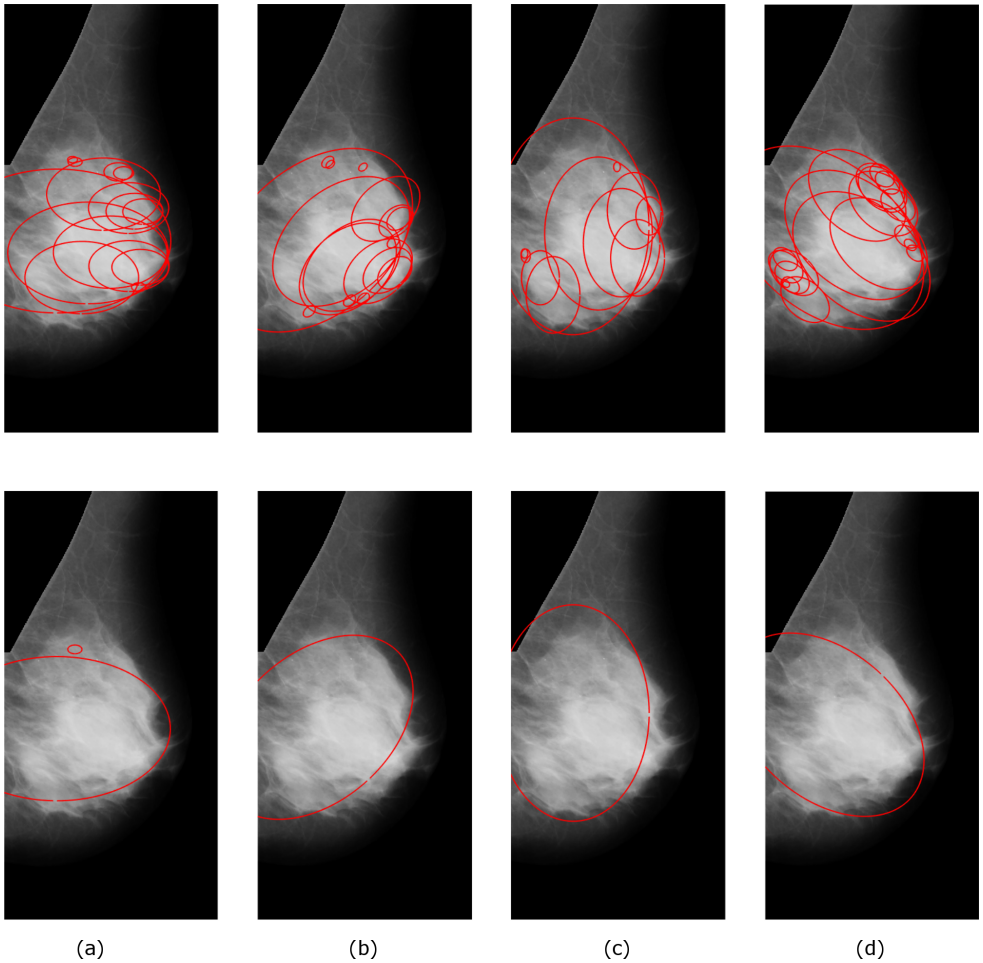



Fig. 4.18 Top row: final blobs at each orientation after calculating the global extremum for different orientation. Bottom row: blobs merged at each orientation.

are partially overlapping as in Figure 4.19 (b), the area of the overlapped region is estimated. Then the overlapped region is reduced from each blob area to estimate the largest blob to retain based on a threshold parameter. Threshold parameter 't' estimates the extent of overlapping and decide whether to include both the blobs or delete the smaller blob. Likewise, if the blobs are completely apart as in Figure 4.19 (c), then both the blobs are retained to represent dense tissue. The blobs merged using area-based merging method is shown in the bottom row of Figure 4.18 which will be later combined along all orientations to retain only the elliptical blobs according to the direction of tissue patterns.

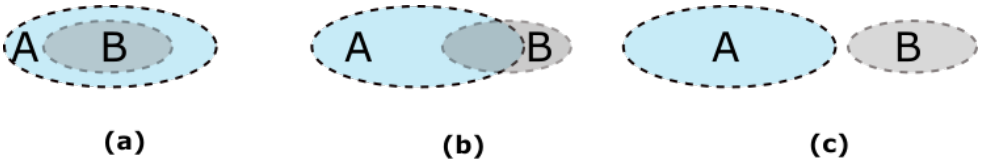



Fig. 4.19 A sample illustration of area-based blob overlapping.

After merging the blobs in each orientation, the retaining blobs from all orientations are combined to compute the largest and most suitably oriented blob relative to breast tissue pattern. The area-based blob merging was performed to combine the oriented blobs. The combined merging of blobs is represented as in Figure 4.20.

Though the blobs are combined at different orientations, the merging with blob area, orientation and distance between blobs finds inappropriate in some cases. When sorted based on blob area for merging, the final orientation 135° will be retained in many cases.

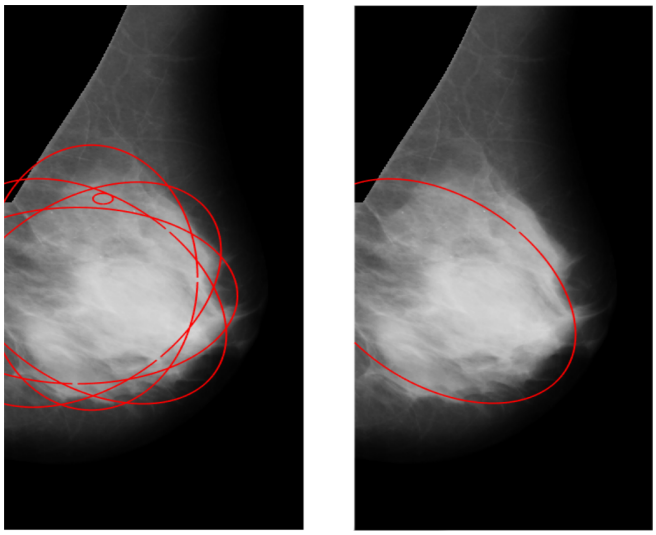


Fig. 4.20 Combined blob merging at various orientations.

### 4.5.3 Results and Discussion

From a visualization perspective and evaluation, it is found that elliptical blob modelling can be used to model the parenchymal breast patterns and in risk assessment. We use the mammogram images from MIAS dataset for the initial study of the proposed model. As discussed initially, the blobs were selected from 10 scales for four different orientations with an initial scale of 8 pixels which increases by a factor of $\sqrt{(2)}$. As per the BI-RADS categorization, BI-RADS-I mammographic breasts almost entirely fatty which means there will be only a few dense tissue patterns, so only a few small scale blobs are detected. For BI-RADS II, the glandular tissue is scattered in a small amount resulting in medium-scale blobs. BI-RADS III have more glandular tissue compared to BI-RADS, resulting in larger blobs. BI-RADS IV category covers the entire breast region with dense tissue resulting in huge blob representing the entire blob region.

To evaluate and visualize the possibility of proposed elliptical blob model, mammographic images from MIAS dataset of BI-RADS category I to IV are shown in Figure 4.21

The merging of blobs tends to be complex due to multiple factors affecting the deletion of blobs like threshold ($t$), orientation, size of blob and number of blobs in each scale. Due to this, most of the merging ends in a blob orientation of 135°. Likewise, $t$ is a factor that decides the amount of accepted overlap between blobs. So, with the variation in $t$, the number of blobs removed from blob pool varies. A sample effect is shown in Figure 4.22. For a mammogram image with 46 overlapping blobs, the number of blobs removed varied with the t parameter. For $t$ = 0.02, the number of blobs retained was 5, while 15 and 41 blobs were retained for $t$=0.7 and $t$=1 respectively.

Though the proposed method seems visually plausible, it needs further improvements to be used for density and risk assessment. According to Bakic's synthetic model, the ducts and patterns could be represented through various sized shells, elliptical blobs. In our approach, we described the tissue patterns in a mammogram in varying elliptical blob scales to show the tissue pattern. We used a generalized model of LoG, but other blob detectors could also be exploited like the Hessian variants, DoG and scale-invariant feature transform (SIFT). In our experiments, we set the initial scale to be eight and increased to 10 scales reaching a scale of 181 pixels. But it will be better to automatically decide the scale than fixing the scale level. Again, with orientation, we selected four directions as an initial study. Still, from experimental studies, the model performance will be better if it is designed to automatically select

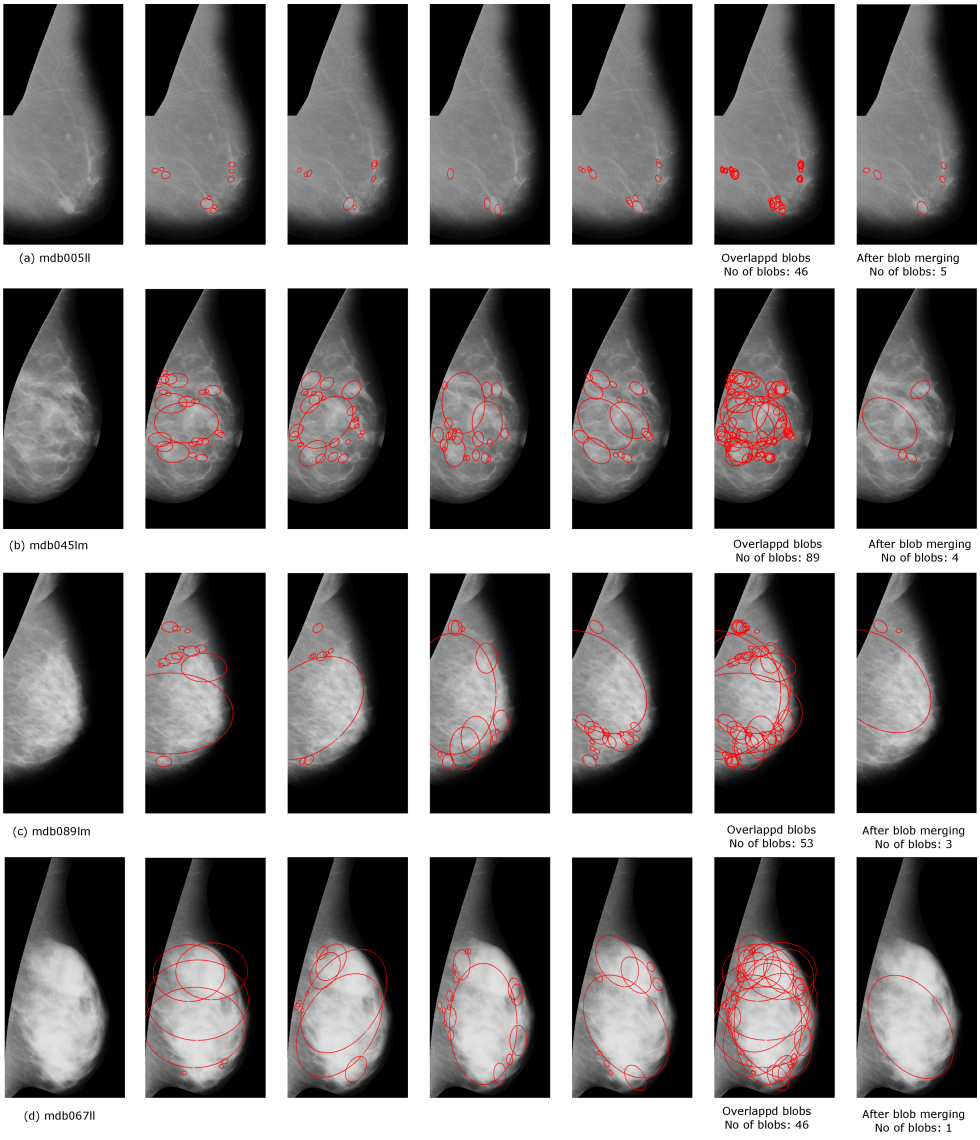



Fig. 4.21 Blob representation with respect to BI-RADS density classes. row 1-4: BIRADS I-IV, col 1: Original images, col 2-5: blobs at orientation 0°,45°,90°,135° respectively, col 6: overlapping blobs, col 7: merged blobs.

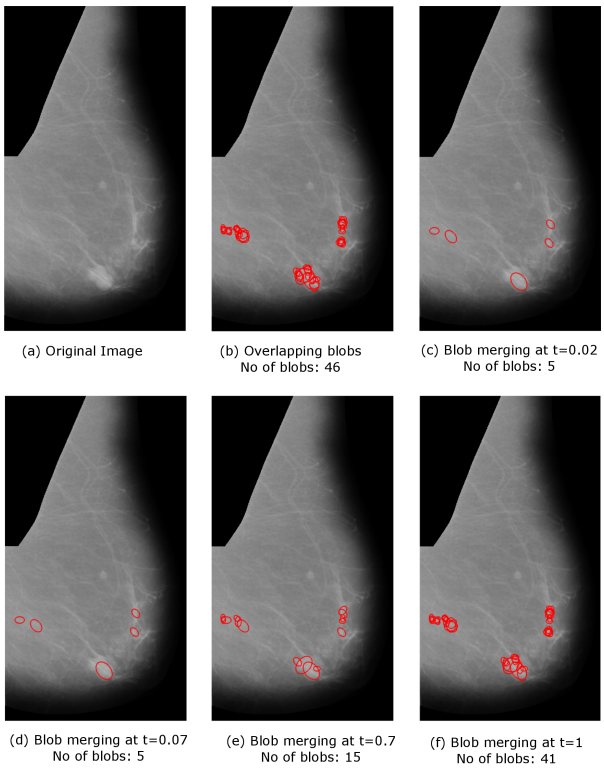



Fig. 4.22 Changes in blob merging depending on the threshold parameter.

the orientation and corresponding scale rather than a fixed direction. Similarly, while considering the merging, more efficient technique that could manage multiple factors can be developed. For example, instead of focusing on orientation, scale level, number of blobs for sorting and merging; an additional factor which estimates the minimum, maximum or mean intensity of blob region along with blob size and direction could be a possibility for better merging.

## 4.6 Summary

We attempt to design an elliptical blob model for representing the parenchymal breast patterns. To the best of our knowledge, this was the first attempt of breast density risk estimation based on fatty and dense tissue pattern correlating with BI-RADS density class. Besides, this was the first attempt to represent breast tissue, analyzing the breast tissue structure orientation using elliptical blobs at multiple scales. For the support of the work, various blob detector models were explained, followed by multiple breast synthetic models. From a visualization perspective and an attempt to explain the tissue type based on elliptical blob patterns on real mammograms, the method found to be successful. The method found to be correlating with BI-RADS density explanation

based on tissue as the amount of area covered by elliptical blob representing dense tissue increases with density class from BI-RADS I to IV. But as future work, it needs further improvements like multi-factor dependent merging to retain the best blob structure to represent the parenchymal patterns similar to Bakic's simulated synthetic models of the breast. Likewise combining circular and elliptical blob structures by exploring there elongation properties, orientation caould be investigated to include more texture patterns and match with the synthetic models generated through blob modelling. The method was evaluated on the MIAS dataset.

# Chapter 5

# Classification of Micro-calcification Clusters using Connected-Chain Modelling

While breast density is considered as a risk factor in developing breast cancer, micro-calcification (MC) are considered as abnormalities that indicate the presence of breast cancer. So, in this chapter, we propose a novel approach to classify the micro-calcification clusters as benign or malignant. The topology/connectivity of individual micro-calcification in a micro-calcification cluster (MCC) is investigated to diagnose the malignancy. The multi-scale morphology of MC in MCC distribution was also investigated to estimate the malignancy. The topology-based models in our proposed approach focus on the distributional pattern, unlike the traditional ways of analysing the morphology and statistical features for classification. The role of surrounding breast tissue is investigated to estimate the role of texture properties in classification. Section 5.3 proposes a new connected-chain model based on MC distributional pattern followed by its multi-scale connected-chain approach in Section 5.4. Finally, the potential of connected-chain model is compared and combined with traditional methods of tissue texture analysis and MC morphology features in Section 5.5.

## 5.1 Introduction

Differentiating the abnormalities/lesions like micro-calcifications in mammograms as benign or malignant is a laborious and time-consuming process for radiologists. As the mammograms are a two-dimensional representation of three-dimensional tissue structure, the overlapping of breast tissue may lead to a pattern similar to an abnormality or

may change the structure of abnormalities resulting in a misclassification (Elter and Horsch, 2009). It is estimated that less than 30% of reported breast biopsies ended up as malignant causing unnecessary stress to patients (Adler and Helvie, 1992; Elmore et al., 2003). But micro-calcifications are considered as a primary indicator of the presence of breast cancer (Sakka et al., 2006). The high correlation of breast cancer and histologically proven MCCs shows the relevance of developing Computer-Aided Diagnosis (CAD) tools in investigating the presence of micro-calcifications in the breast. Although CAD systems can assist the radiologists in early detection of many abnormalities as a reference, the interpretation/classification of micro-calcifications remains difficult and challenging due to the fuzzy nature of micro-calcification, the small size, although having high contrast there is potential overlap with dense tissue making it difficult to distinguish from surrounding tissue, absence of particular patterns or templates due to the variation in sizes, shapes and distribution (Cheng et al., 2003; Sakka et al., 2006). It is estimated that the size of micro-calcifications ranges from 0.1 mm to 1.0 mm, with an average size of 0.3 mm. Similarly, due to high-frequency noise and overlapping patterns, some isolated micro-calcifications are challenging to be distinguished in the film-screen mammography (Sakka et al., 2006). So, enhancement and noise filtering can be beneficial for the localisation of micro-calcifications in breast tissue (Cheng et al., 2003).

Similarly, extracting distinguishable features to classify micro-calcifications are a vital characteristic of a CAD system. The features extracted should be reliable, independent (uncorrelated), discriminative for efficient classification. Features can be extracted and distinguished based on the extraction method and image properties such as statistical, shape features, topological graph features, cluster feature extraction which are discussed in detailed in Section 2.4.2.

## 5.2 Related Work

Various features have been used to classify micro-calcifications, for example, individual micro-calcification characteristics describing the perimeter, area, compactness, elongation, eccentricity, thickness, orientation, and contrast (Sakka et al., 2006). Other features used for classification include co-occurrence features, wavelet properties, Gabor filter bank features, scale-space characteristics, fractal and cluster features (Cheng et al., 2003; Elter and Horsch, 2009) (see Section 2.4.2 for detailed feature extraction and classification methods). Elter and Horsch (2009) described shape, size, cluster features, intensity, texture, morphology and topology features for classification of

micro-calcifications. Some of the wavelet features (Docusse et al., 2008) used for classification included energy and entropy while the cluster features for classification described the distribution of local features, cluster area, cluster perimeter, cluster diameter, eccentricity, elongation and number of micro-calcifications in the cluster representing the cluster morphology (Betal et al., 1997).

According to breast micro-calcifications distribution studies, it is found that malignant micro-calcifications tend to be small in size and densely distributed (>5 per focus within $1cm^2$), while benign micro-calcifications are usually larger in size, smaller in number and scattered (<4-5 per $1cm^2$) (Feig et al., 1987; Nalawade, 2009; Sickles, 1986). Numerous methods were developed to explore the micro-calcification cluster features for classification of abnormality as benign or malignant. Veldkamp et al. (2001) investigated the cluster shape and distribution features like the number of calcifications in the cluster, mean and standard deviation of micro-calcification area, orientation, contrast, shape and, position of a cluster in the mammogram. Similarly, Foggia et al. (2001) investigated the shape, distribution of micro-calcifications within the cluster and irregularity of micro-calcifications were explored, assuming that more irregular and non-uniform shapes of micro-calcifications represent a malignant cluster. Winkler et al. (2000) utilized the features and patterns of micro-calcifications inside a cluster, especially the number of micro-calcifications in the cluster, cluster area, perimeter, and the standard deviation of inter micro-calcification distance within the cluster. Shao et al. (2011) investigated the cluster ROI features like the population density of dispersed MCs, the smallest diameter to encircle the clustered region, diameter of largest micro-calcification, diameter of smallest MC in the ROI and mean diameter ensemble average for all MCs in the ROI. Estimating the distribution and closeness of micro-calcification in the clusters, Chen et al. (2015) proposed a topology-based classification method by constructing graphs at multiple scales. The graph features were estimated at different scales for the topological modelling and classification of micro-calcification into benign and malignant cases. Similarly, Suhail et al. (2018b) used a tree-based topological approach for the classification of micro-calcification, focusing on the distribution and connectivity of micro-calcification. The tree features like the height of the tree, and the number of leaf nodes were evaluated for the classification process.

## 5.3 Connected-Chain Model for Micro-calcification Classification

Mammography can identify some abnormalities before they are detectable physically. Breast cancer can appear in mammograms as abnormalities like masses, architectural distortions or micro-calcifications. While breast masses appear as large white bright regions in mammograms, micro-calcifications appear as bright, small in size and form a cluster. The radiologists investigate the presence of fine, patterned granular micro-calcification clusters as an indication of early breast carcinoma and perform further histological examinations for final confirmation. To develop a CAD system for diagnosis of micro-calcifications, researchers evaluate the calcifications based on their primary properties like size, shape, form, distribution, density and distribution as described in Section 5.2.

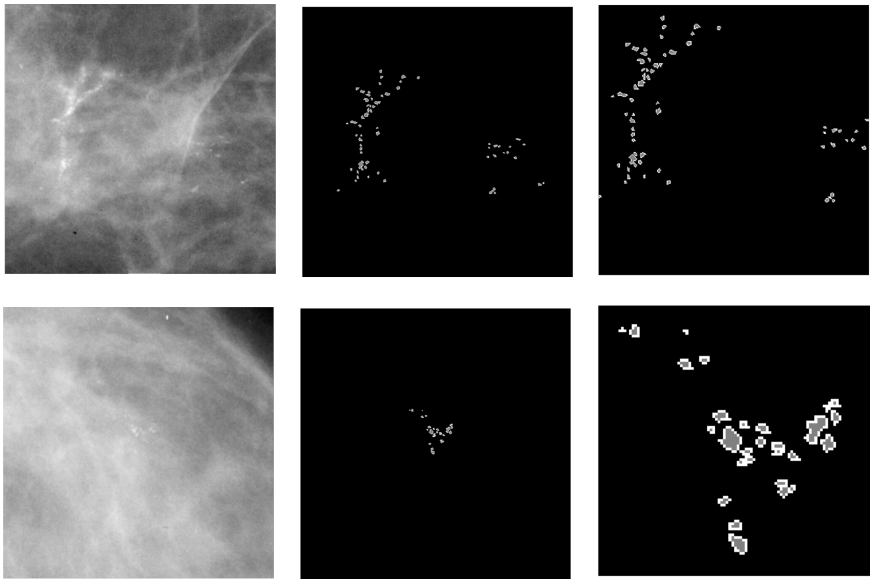


Fig. 5.1 Examples of ROI patches of malignant (top row) and benign (bottom row) mammographic micro-calcification clusters; First column: original patch; second column: annotated micro-calcifications; third column: zoomed in view of micro-calcifications.

The proposed approach deals with the classification of manually or automatically detected/segmented ROIs of micro-calcification clusters as benign or malignant using topological features. Unlike the tree-based approach by Suhail et al. (2018b) and topology-based multi-scale graph model by Chen et al. (2015), the proposed method uses the closeness and distribution of micro-calcifications in local clusters, investigates the distribution pattern of micro-calcifications in the clusters, which is linked to the

clinical aspects in estimating the classification of micro-calcifications (Hernández et al., 2016; Nalawade, 2009; Park et al., 2000).

### 5.3.1 Data and Methodology

The proposed approach used the clinical perspective of micro-calcification distribution in a cluster that the benign micro-calcifications tend to be of larger size and more widely spread compared to malignant cases with relatively smaller size and closely distributed (Hernández et al., 2016).

#### Dataset

The dataset used for initial evaluation of the proposed algorithm was segmented ROIs of mammogram images from the DDSM database. A total of 289 mammogram ROIs of varied sizes with micro-calcification abnormalities were evaluated, where 131 mammogram ROIs were histologically proven as malignant cases while 158 mammogram ROIs were histologically proven as benign cases. The average size of these ROIs was $482 \times 450$ pixels, though it should be mentioned that the proposed method is independent of the patch size. (see Section 2.9 for more details on the dataset).

#### Methodology

In the proposed approach, the micro-calcification ROI clusters detected/segmented by Oliver et al. (2012) automatic detection method was used for feature extraction and classification. The detection method extracts the local features that characterize the morphology of micro-calcification using a bank of filters. Afterwards, a boosting scheme is used to obtain the most salient features to detect individual micro-calcifications. Later, the probability images of individual micro-calcifications are calculated to determine the cluster.

Subsequently, the detected mammogram ROIs are binarized by converting them to binary images. All the pixels with value '0' represented background or healthy tissue while the pixels with value '1' indicated micro-calcifications. Single and low probability pixels were removed as they were considered noise. So, the denoising was performed by eliminating low probability pixels and by deleting small areas with a size smaller than 4 pixels as shown in Figure 5.2, which demonstrates the initial steps of automatic localization, binarization and denoising of image ROIs. A morphological erosion operation was performed to remove the micro-calcifications of size less than

four pixels. Later, a dilation process was performed with the same structuring element to recover the original size of the remaining micro-calcifications.

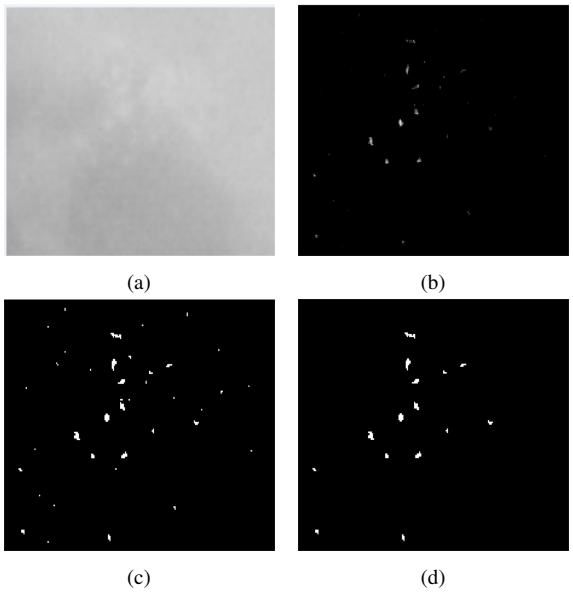



Fig. 5.2 Binarization and denoising of annotated micro-calcification cluster; (a) original mammographic patch with micro-calcification, (b) annotated micro-calcification cluster image, (c) binarized image, (d) denoised image.

After pre-processing, the centroids of each micro-calcifications were estimated. After that, the distance between each centroid point was calculated to determine their relative position. Connected-chains were constructed for determining the closeness of MCs. Connected-chains were estimated using a threshold distance of 40 pixels to estimate the longest possible connected-chain. The connected node chains were being built by assessing and joining the nodes which are distributed within 40 pixels around each node (the Euclidean distance measure (Lele and Richtsmeier, 1991) was used for distance calculations). The first node was taken, and the next closest node to it from the distance map was joined with it, followed by the nearest node to the previously connected node. The procedure was continued until there were no unconnected close nodes to any of the nodes in the constructed chain. The procedure was repeated by selecting the next unvisited node from the node list to start the next chain. So, each connected-chain represented a cluster of micro-calcifications which are jointly distributed with each other. The overview of the proposed method is illustrated in Figure 5.3. The connected nodes in the chain show the pattern/closeness of node arrangement in that cluster, which was the representation of closely located connected nodes. In other words, the number of connected-chains represents the independent number of chain-cluster. Those micro-calcifications which are closely

located will form a chain-cluster (3,4,5,7,8,9,10,11,12,13,14,15), (1,2) while scattered micro-calcifications are considered as unvisited or leaf nodes (6) as shown in Figure 5.4. Figure 5.5 illustrates how the connected-chains are generated from the distance map (the nearby centroids/nodes are selected uniquely for each visited node/centroid (3,4,5,10,11,7,8,9,12,13,14,15) and sorted (3,4,5,7,8,9,10,11,12,13,14,15) to form a chain). In both cases, the connected-chains formed contains same nodes though it shows different pattern of representation.

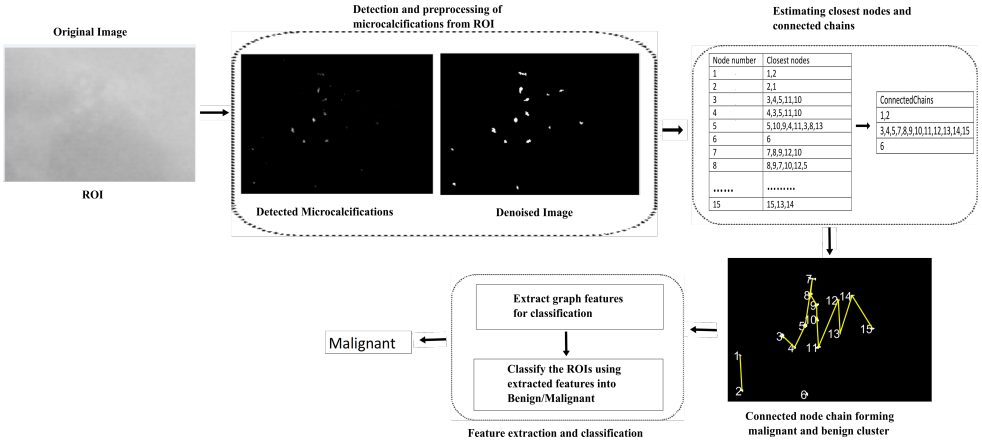
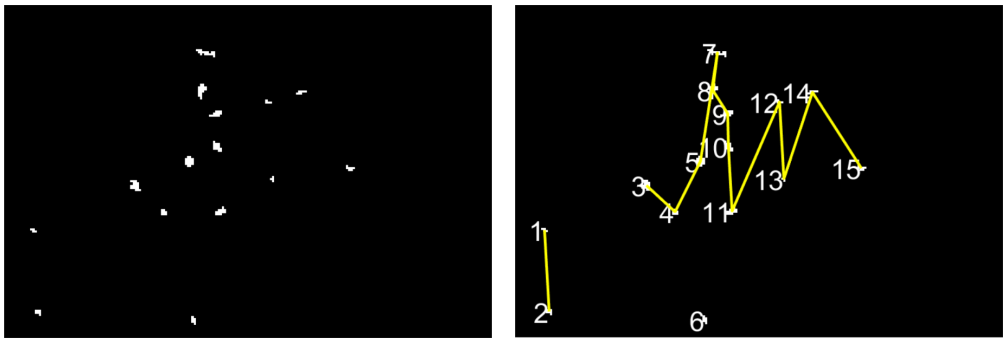



Fig. 5.3 Detailed representation of the proposed connected-chain graph method for micro-calcification classification into benign/malignant.



(a) (b)

Fig. 5.4 Generating connected-chains for micro-calcifications to construct clusters; (a) the micro-calcifications after denoising, (b) representing the cluster/connected node chain, which indicates three chain-clusters of various sizes.

## Feature Extraction

Cluster properties are an important characteristic feature of micro-calcifications in detecting its type of abnormality as benign or malignant. If micro-calcifications are

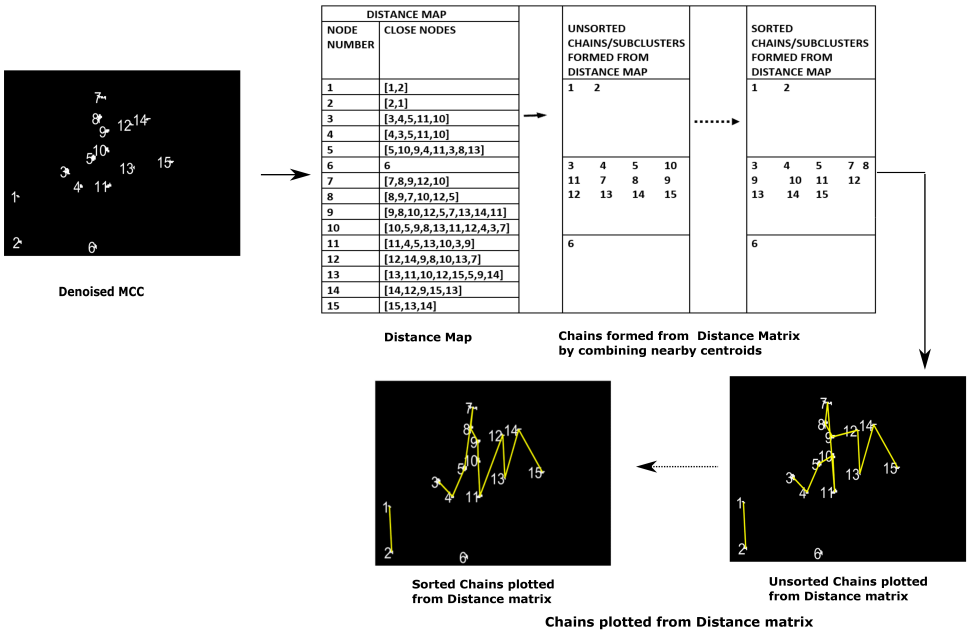



Fig. 5.5 Detailed representation of generating connected-chains from the distance map.

scattered throughout the breast, it indicates a benign abnormality (Oliver et al., 2012). So, in the proposed approach, we extract the features of connected-chain, considering it as a graph. For feature extraction from the chain-clusters, the cluster properties calculated were the number of clusters (independent number of connected-chains), the number of benign clusters (the chains containing less than or equal to 5 nodes), the number of malignant clusters (the chains with more than five nodes), the size of the longest chain, the number of independent nodes/leaf nodes. We take the minimum chain length to be five nodes as a threshold chain length. Micro-calcifications are considered as a cluster when there are more than three micro-calcifications in a $1cm^2$ square area (which is equivalent to $200 \times 200$ pixels in the digitized data (DDSM and MIAS) with a pixel size equal to 50m) (Alam et al., 2019; Kopans 3rd, 2007; Oliver et al., 2012). Similarly, from a clinical perspective, the malignant clusters are densely packed and are fine and tightly structured. So, we took a connected-chain length threshold as more than 5 for the malignant chains. Therefore, in our study, we investigate the MCs based on how densely/scattered they are distributed in a MCC in classifying a MCC as benign/malignant.

From a clinical perception, the malignant chains will be more complex and tightly packed compared to benign clusters. Figure 5.6 illustrates the difference between connected chains formed for benign and malignant cases. The first column represents the mammographic patch while the second column shows the segmented and

histologically proven calcifications and the third column represent the connected-chains for benign and malignant cases. The connected-chains structure for malignant micro-calcifications tend to be more complex compared to the benign structure as the malignant micro-calcifications are tightly and densely arranged, giving a long, dense chain.

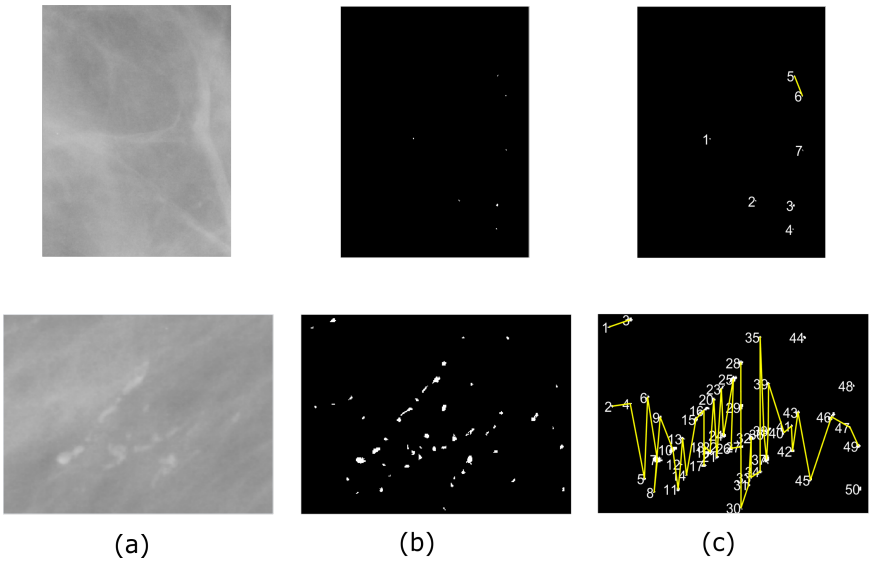



Fig. 5.6 Complexity representation of benign (top row) and malignant (bottom row) example ROIs; columns:-(a): original mammogram ROI; (b): the detected binarized micro-calcification; (c): connected-node chain for benign and malignant clusters.

### Experimental Results and Discussion

The efficiency of features extracted from connected-chain spatial distributional model for classification of mammographic ROI's into benign and malignant cases were evaluated using various classifiers like the Bayesian Network (Sharma et al., 2008), the k-nearest neighbour (KNN) classifier (Sharma et al., 2008), AdaBoost classifier (Friedman et al., 2000), SVM (Efros and Freeman, 2001), and Random Forest (Taylor et al., 1999). The classical KNN classifier is an instance-based learning approach. It is based on simple majority voting unless equal class probability is indicated, and the Euclidean weighted approach is used as the distance measure. Bayesian network uses a directed acyclic graphs (DAGs) to represent the joint probability distribution of set of random variables where edges represent direct correlations. A simple estimator is used for estimating the conditional probability tables and a hill-climbing algorithm was used as a learning algorithm. In order to estimate the performance of the method, the features extracted are classified into malignant or benign cases using a 10-run 10-FCV and leave-one-out approach using the mentioned classifiers.

Table 5.1 Comparison of classification accuracy using different classifiers.

| | Automatic Classification | |
|---|---|---|
| | 10 FCV | Leave-one-out |
| Bayesian Network | 84.602±0.496 | 84.78 |
| AdaBoostM1 | **85.603±0.333** | **85.81** |
| SVM | 83.183 ±0.467 | 83.04 |
| Random Forest | 84.320 ±0.879 | 83.74 |
| KNN | 84.010±1.127 | 84.08 |

Table 5.2 Area under ROC ($A_z$) for different classifiers.

| | Automatic Classification | |
|---|---|---|
| | 10 FCV | Leave-one-out |
| Bayesian Network | $0.839 \pm 0.010$ | 0.844 |
| AdaBoostM1 | $0.839 \pm 0.010$ | 0.779 |
| SVM | $0.817 \pm 0.005$ | 0.816 |
| Random Forest | **0.912±0.006** | **0.916** |
| KNN | $0.888 \pm 0.004$ | 0.891 |

The classification accuracy by various classifiers are mentioned in Table 5.1. The results from the table shows that all the classifiers performed similar in classifying the features into benign/malignant with slight variation due the classifier feature selection methods used in classifiers, with AdaBoostM1 obtaining the best CA% result of $85.603 \pm 0.333$ for the extracted features. The Bayesian network, Random Forest and KNN attained similar CA% results of $84.602 \pm 0.496$, $84.320 \pm 0.879$, $84.010 \pm 1.127$, respectively. The SVM classifier performed poor compared to others with CA% of $83.183 \pm 0.467$. In addition, the area under ROC curve ($A_z$) is calculated as in Table 5.2 obtaining $0.912 \pm 0.004$ for the Random Forest classifier. The $A_z$ results shows that Random Forest performed better compared to other classifiers in the list. While considering the $A_z$ performance of listed classifiers, KNN obtained $0.888 \pm 0.004$ followed by similar perfomance by Bayesian Network and AdaBoostM1 with $0.839 \pm 0.010$. Comparing to other classifiers in the list, SVM classifier performed poor with $A_z$ of $0.817 \pm 0.005$ for the extracted features.

Though the proposed approach can successfully classify most of the cases correctly, there are misclassified instances like the ROIs shown in Figure 5.7. The top row ROI was misclassified as a benign case by the proposed approach where it only has one connected-chain with less than 6 nodes and the other nodes are scattered giving less probability of malignancy. The second row was misclassified as malignant because of the large number of connected chains though the complexity of each connected chain

is low giving an assumption of malignancy even when the micro-calcifications were spread out over the patch.

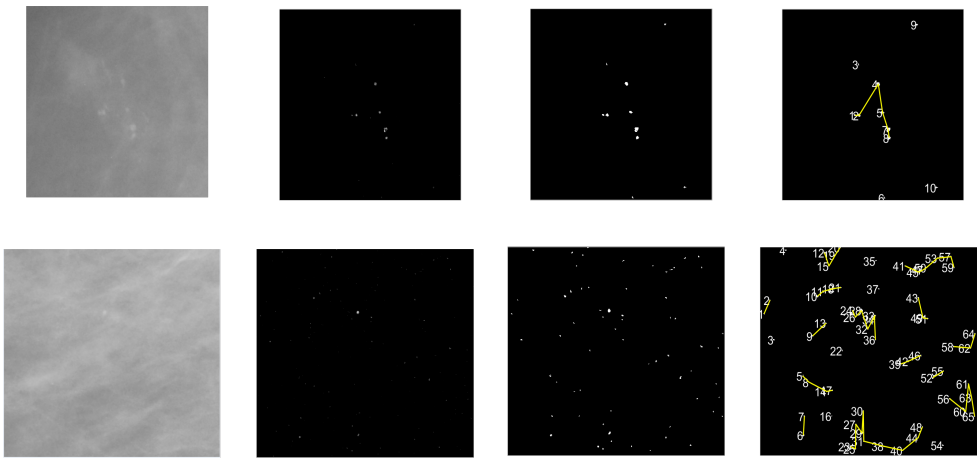


Fig. 5.7 Misclassified examples. First row: a malignant ROI reported as benign; Second row: benign ROI reported as malignant.

The proposed method utilizes the closeness and number of micro-calcifications that are arranged tightly in a cluster. To evaluate the connected-chain pattern and investigate the orientation dependence, we flip the image orientation and find the connected-chain. From the experimental evaluation, it was found that the cluster-chain is similar in both orientations. A few sample images are provided in Figure 5.8. The first column shows the original micro-calcification cluster ROIs; the second column shows the denoised micro-calcification cluster, the third column shows the connected-chain pattern while the last column shows the pattern while the ROI patch is oriented. It is to be noted that the chain-cluster formed in both cases are similar to the same set of micro-calcifications in the chain.

The efficiency of the proposed approach is comparable with the state of art methods for micro-calcification classification. It will be discussed and compared in detail in the next section as most of the topological methods in the literature utilizes the multi-scale approaches of graph and tree structures.

In the discussed method, the connectivity features were utilized only from a single scale. So to evaluate the effect of image resolution on classification, multi-scale connectivity features are to be considered, and more datasets are to be assessed, which helps in evaluating the robustness of the proposed approach. The multi-scale connected-chain approach on various databases and a comparative study with other state of art methods for micro-calcification classifications will be discussed in detail in Section 5.4.

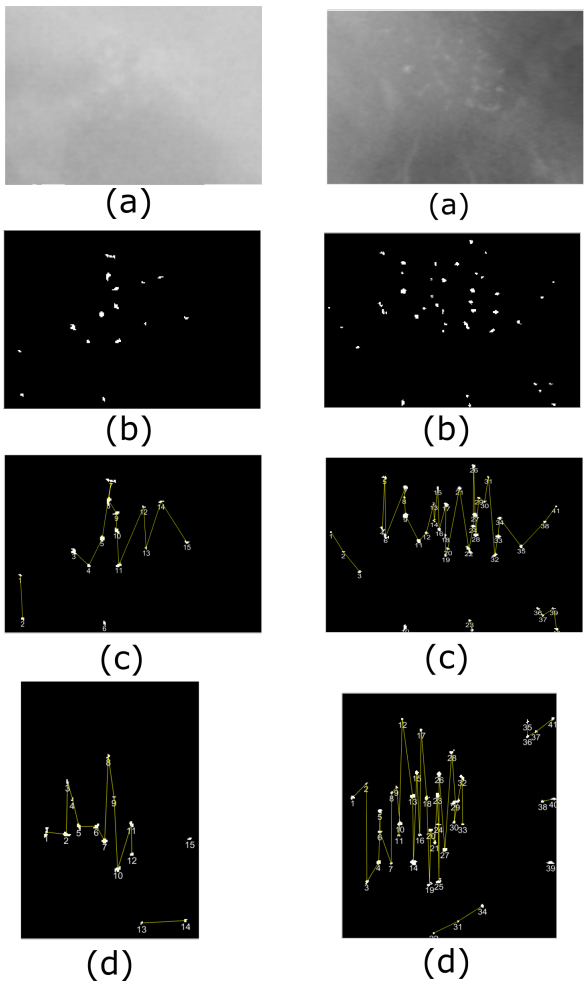



Fig. 5.8 Connected-chain clusters formed for oriented ROIs: rows:- (a) original ROIs, (b) denoised binary image, (c) chain pattern for original ROI, (d) chain pattern for flipped ROI.

## 5.4 Multi-scale Connected-Chain Model for Micro-Calcification Classification

To explore the topological and multi-scale distributional connectivity features of micro-calcifications, the connected-chain method was modified to extract features from multiple scales/levels. The proposed approach used the multi-scale morphological relationship of connectivity between micro-calcifications where connected-chains between nearest micro-calcifications were generated at each level. Subsequently, graph/chain connectivity features at each scale were extracted to estimate the topological connectivity structure of micro-calcification clusters for benign versus malignant classification. The proposed approach was evaluated using publicly available digitized datasets: MIAS and DDSM in addition to the digital OPTIMAM dataset as most of the related

works are based on these datasets. The study showed that topological/connectivity modelling using a multi-scale approach was appropriate for micro-calcification cluster analysis and classification. The topological connectivity and distribution can be linked to the clinical understanding of micro-calcification arrangements (Park et al., 2000).

### 5.4.1 Data

A multi-scale approach was developed, and the algorithm was performed on different datasets to make the method independent of resolution effect (see Section 2.9 for more information on the datasets). The DDSM and MIAS database were used for analysis from which relevant MCC detected regions of interest (ROIs) were extracted for evaluation. As discussed, 289 mammogram ROIs of varied sizes (average size of 482×450 pixels) from DDSM that were histologically confirmed and automatically detected/segmented were used. Similarly, there were in total of 20 images in the MIAS dataset containing micro-calcification cluster ROIs (image patch size of 512 ×512 pixels) with 11 benign clusters and nine malignant cases which were all categorised by histology.

In addition to DDSM and MIAS, the digital dataset, OPTIMAM was used. A total of 286 mammographic images with micro-calcification clusters were used for estimating the robustness of the algorithm, where 136 ROIs were histologically reported as benign and 150 were categorised as malignant. The micro-calcifications for OPTIMAM were segmented using the detection approach developed by Alam et al. (2018). This was done to study the segmentation and resolution effect of mammograms on classification.

### 5.4.2 Proposed Methodology :- Multi-scale Connected-Chain Modelling

The proposed methodology used the micro-calcification abnormality patches as the input which were binarized and denoised for the classification process. The closeness of micro-calcifications was estimated through connected-chain graphs at different scales. Then, the extraction and classification of connected-chain graph features at respective scales was used for classification as malignant or benign. A detailed description showing each step of the process is shown in Figure 5.9 and a brief explanation can be found in of Algorithm 1.

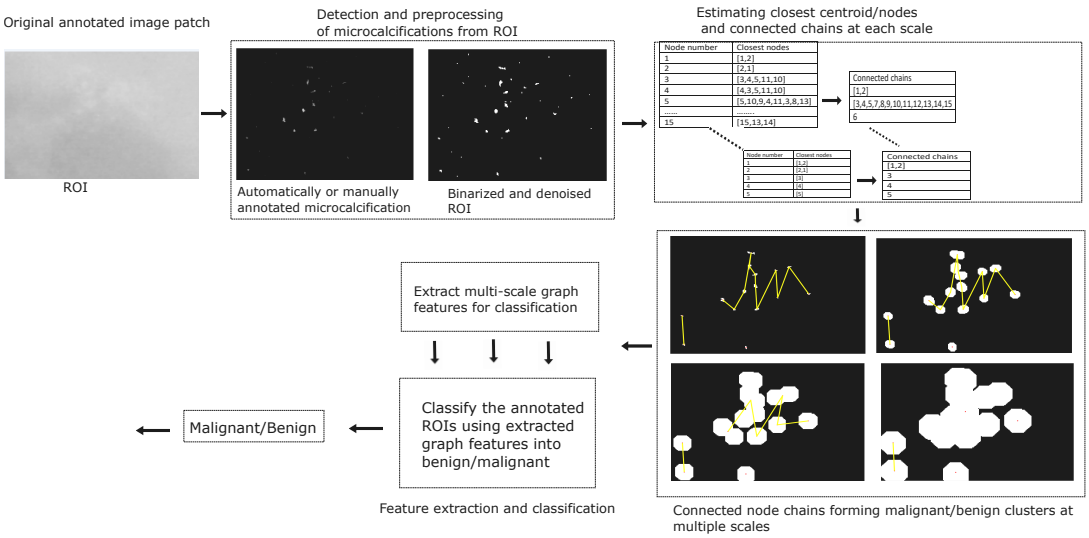



Fig. 5.9 Detailed representation of the proposed multi-scale connected chain graph method for benign/malignant micro-calcification classification.

### Pre-processing

The automatically segmented mammographic ROIs are binarized for further morphological operations. Denoising is performed on the binary image by removing regions which are less than 4 pixel considering those as noise or as low probability micro-calcification regions as in the previous study. The binarization and pre-processing procedures are demonstrated in detail in Figure 5.2.

### Constructing Connected-Chains at Multiple Scales/levels

Following denoising, the centroid of each micro-calcification was calculated for finding the nearest micro-calcifications. Initially, all centroids were considered as independent nodes. The number of scales/levels was set (e.g. $S = 3$, i.e. the number of times the morphological dilation will be performed). We selected $S$ to be equal to three because a large scale number with large structuring element for dilation can morphologically merge all the micro-calcifications at an early stage, making it a single unit. At each scale, a morphological dilation operation using a disc-shaped structuring element with 5 pixels radius size was performed for estimating the connectivity between micro-calcification centroids. Similarly, reducing the size of the structuring element would need additional scales to merge the micro-calcifications.

The distance between each centroid point was calculated to estimate the closeness in distribution generated by each distance map. Subsequently, connected-chains were created to describe the morphological distribution of each micro-calcification

with respect to other micro-calcifications. For the first scale, nodes were connected, which were distributed within 40 pixels. The Euclidean distance measure (Lele and Richtsmeier, 1991) was used for distance calculations. The initial node from the distance map was selected as the first node in the chain. The procedure continued by selecting the next unvisited node to start the next chain. The method repeated recursively until all the nodes were visited. So, each connected chain represented a sub-cluster of a micro-calcification cluster. For the following scale, morphological dilation (Soille, 2013) using a disk structural element of size five was performed on the binary micro-calcifications and the centroids and distance map were calculated followed by the connected-chain estimation.

The connected-chains at each scale/level represented the arrangement or pattern of micro-calcification in the clusters. The connected-chain structure differed with each level as some dilated micro-calcifications merged, giving a single centroid for distance estimation. The closely distributed micro-calcifications merged at the initial stages of dilation forming a single unit. Therefore, the morphological dilation process gave different connectivity patterns between the micro-calcifications, building different independent sub-clusters. The detailed multi-scale space clustering is illustrated in Figure 5.10. Since the malignant micro-calcifications were closely located, with increasing scales, they tended to merge early as a region. In contrast, the benign micro-calcifications needed additional levels to be united as they were more diffusely distributed.

**Feature Extraction** In the connected-chain each node is the centroid of a micro-calcification and is considered independent. It is noted that the number of connections between nodes decreased with increasing scales as the micro-calcifications started merging. The graph features of the connected-chains at each scale were extracted and concatenated to form the feature vector for classification. The cluster properties calculated at each scale were the number of clusters, the benign clusters (the chains containing less than 6 nodes), the number of malignant chains (the chains with more than 5 nodes), the size of the longest chain, the number of independent nodes/leaf nodes, similar to the previous study.

**Classification** The classification of mammographic patches into benign or malignant cases using the connected-chain features at different scales was performed using classical k-nearest neighbour (KNN) (Duda et al., 2012) on the MIAS, DDSM and OPTIMAM databases. We selected the KNN classifier as most of the literature used

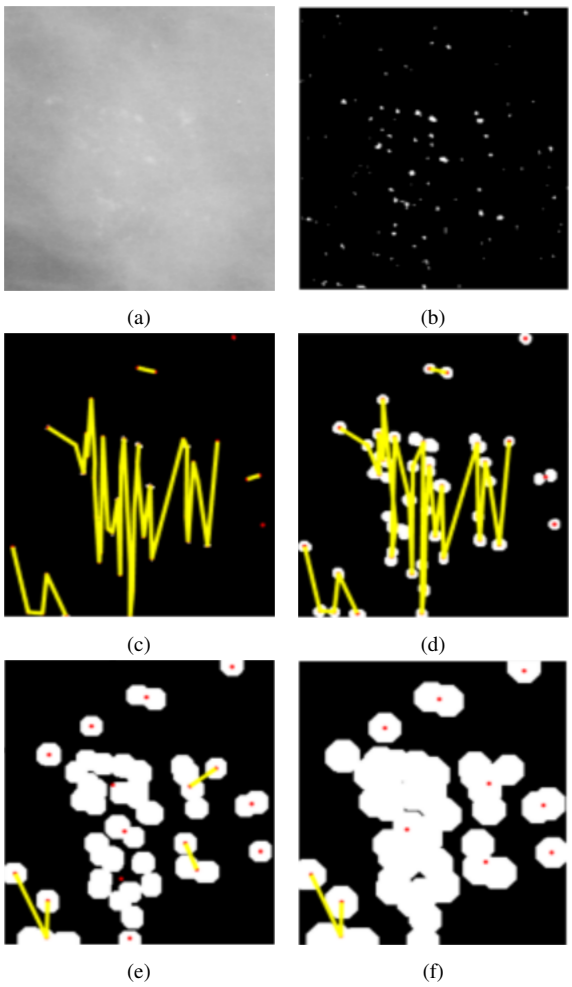



Fig. 5.10 Connectivity pattern of micro-calcifications with morphological dilation at increasing scales:, (a) original image ROI, (b) segmented micro-calcification mammographic ROI, (c) connected-chains at scale 1, (d) connected-chains at scale 2 (micro-calcifications have started merging), (e) connected-chains at scale 3 (most of micro-calcifications have merged), (f) connected-chain at scale 4 (majority of micro-calcifications overlapped forming a single unit).

the classical KNN approach although it should be clear that alternative classifiers are possible as performed in the previous study.

### Experimental Results and Discussion

The graph features extracted at different scales were used with a leave-one-out, and ten-fold cross-validation (FCV) approaches on the MIAS, DDSM and OPTIMAM datasets to investigate how significantly these approaches affected the classification accuracy. We used the KNN classifier with a ten-run 10-fold cross-validation scheme to calculate the accuracy of classification on a different dataset. The 'k' value (the number of

neighbors used for classification) was selected based on the best classification accuracy (a small value of k means that noise will have a higher influence on the result and a large value of k make it computationally expensive). An example of the variation of classification accuracy with k values for the DDSM dataset (highest CA% when k=3) is shown in Figure 5.11. The same procedure was applied for the MIAS and OPTIMAM datasets for selecting the k value for the KNN classifier. The best classification results using the KNN (k=3) classifier on the DDSM dataset is illustrated in Table 5.3, which shows an average classification accuracy (CA%) of 86.47 % ± 5.94 % for a 10 run 10 fold cross-validation and 87.5% for the leave-one-out approach (see Table 5.6). The sensitivity/recall found to be 86.76%, precision of 86.86% with an F-score (to measure the performance of the test for the positive class) of 86.34% for the DDSM dataset.

Similarly, the classification of micro-calcifications as benign or malignant for the MIAS dataset using the KNN classifier (k=3, see Figure 5.11) gave a classification accuracy (CA%) of 82.5 % ± 2.63 % for 10 run 10-FCV and 80.0% for leave-one-out

**Algorithm 1:** Multi-scale connected-chain algorithm

**Input:** Automatically or manually segmented micro-calcification cluster ROIs, threshold distance, number of scales/levels.
**Output:** Multi-scale connected-chain features.

```
Binarize and denoise the image to remove low probability micro-calcification
  regions from ROI.
Compute the centroid of each micro-calcification considering them as
  independent nodes of the chain.
for scale less than 4 do
    Generate a distance map by computing the closest nodes to each node
    based on the threshold distance.
    for scan each node in the distance map node-list do
        Merge the closest node to each node in distance map to form a chain
        till all nodes in the node list are visited.
    end
    Extract the chain features from each scale like number of independent
      nodes in each scale, number of malignant chains (chains with more than 5
      nodes), number of benign nodes (chains with less than 6 nodes) and the
      size of longest chain.
    Increment the scale/level number
    Dilate the micro-calcification objects in ROIs using a disc structuring
      element of size 5 pixels.
end
Generate the feature vector for classification by merging the features extracted
  at different scales.
```

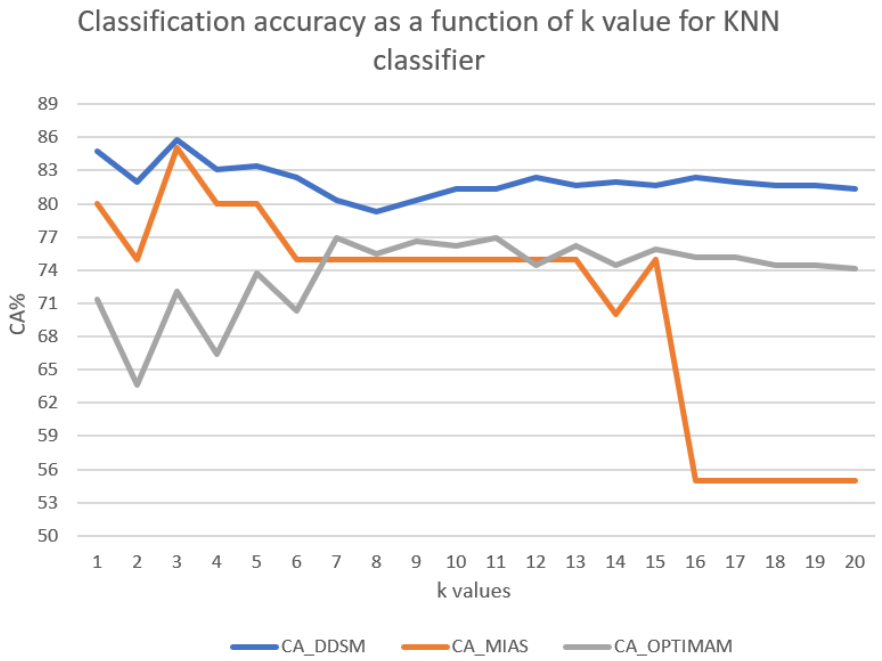



Fig. 5.11 Variation of classification accuracy (CA%) with k-values for KNN classifier (The dotted line shows the trendline of classification accuracy with-respect to k values).

approaches (see Table 5.4 (for the best classification accuracy) and Table 5.6). The test results shows a specificity of 82.25%, sensitivity of 82.50%, and F-score of 83.05%.

Subsequently, the classification of micro-calcifications as benign or malignant for the OPTIMAM dataset using the KNN classifier (k=11, see Figure 5.11) gave a classification accuracy (CA%) of 76.75 % ± 0.66 % and 77.27% for 10-run 10-FCV and leave-one-out approaches (see Table 5.6), respectively. The best classification accuracy for the 10 fold cross validation is illustrated in Table 5.5. The specificity and sensitivity/recall for the classification test for OPTIMAM using the KNN classifiers was found to be 76.68%, 76.69% respectively, attaining the F-score of 76.63%.

The second evaluation process used for investigating the efficiency of classification was performed by ROC (Receiver Operating Characteristic) under the curve analysis. The area under the ROC curve ($A_z$) for the three datasets is shown in Table 5.7 for the 10-run 10-FCV.

As DDSM and MIAS datasets have similar resolution (50 micron), we combine the datasets (309 cases) for MCs classification. We performed a 10-run 10-FCV and leave-one-out classification using KNN classifier (k=3). The classification obtained a CA% of $84.142 \pm 1.22$ with $A_z$, specificity, sensitivity and F-measure of $0.880 \pm 0.010$, 0.85, 0.84, and 0.84 respectively. The leave-one-out approach obtained a CA% of

Table 5.3 Confusion matrices for automatic classification using a KNN classifier for 10-FCV (CA=87.88%) and Leave-one-out (CA=87.54%) approaches for DDSM.

| | | Automatic Classification | | | |
|---|---|---|---|---|---|
| | | 10 -FCV | | Leave-one-out | |
| | | **Benign** | **Malignant** | **Benign** | **Malignant** |
| **Truth Data** | **Benign** | 149 | 9 | 148 | 10 |
| | **Malignant** | 26 | 105 | 26 | 105 |

Table 5.4 Confusion matrices for automatic classification using a KNN classifier for 10-FCV (CA=85.0%) and Leave-one-out (CA=80.0%) approach for MIAS.

| | | Automatic Classification | | | |
|---|---|---|---|---|---|
| | | 10 -FCV | | Leave-one-out | |
| | | **Benign** | **Malignant** | **Benign** | **Malignant** |
| **Truth Data** | **Benign** | 10 | 1 | 10 | 1 |
| | **Malignant** | 2 | 7 | 3 | 6 |

84.79% with $A_z$, specificity, sensitivity and F-measure of 0.886, 0.85, 0.85, and 0.85 respectively.

While performing a quantitative comparison with alternative state of art methods which used topological methods for micro-calcification classification using the DDSM dataset, our results are comparable with the results obtained by Chen et al. (2015) using multi-scale graph modelling with a classification accuracy (CA%) of 85.2 % ± 5.7 % for a set of 300 ROIs without feature selection and Suhail et al. (2018b) with classification accuracy of 91.0% using tree-based modelling on a subset of 129 ROIs from the DDSM dataset. The classification accuracy of our approach was found to be 86.47 % ± 1.30 % for a set of 289 ROIs from the DDSM dataset. Additional comparison of our results with those achieved by other related work with the same datasets are shown in Table 5.8.

While comparing to other topological modelling methods in detail for micro-calcification classifications on the DDSM dataset using the KNN classifier, Strange et al. (2014) attained classification accuracy of 80.0% for a set of 300 cases and modelled the relationship between micro-calcification regions in the form of mereotopological barcodes. Similarly, Chen et al. (2015) utilized morphology and the spatial connectivity relationship for building a multi-scale graph model where two nodes are linked as edge if the corresponding micro-calcifications overlap each other. The

Table 5.5 Confusion matrices for automatic classification using a KNN classifier for 10-FCV (CA=77.27%) and Leave-one-out (CA=77.27%) approach for OPTIMAM.

| | | Automatic Classification | | | |
|---|---|---|---|---|---|
| | | 10-FCV | | leave-one-out | |
| | | **Benign** | **Malignant** | **Benign** | **Malignant** |
| **Truth Data** | **Benign** | 99 | 37 | 99 | 37 |
| | **Malignant** | 28 | 122 | 28 | 122 |

Table 5.6 The classification accuracy for MIAS, DDSM and OPTIMAM datasets for 10-run 10-FCV and Leave-one-out approaches.

| | | CA | |
|---|---|---|---|
| | | **10-FCV** | **Leave-one-out** |
| **Dataset** | **MIAS** | 82.50 % ± 2.63 % | 80.0% |
| | **DDSM** | 86.47 % ± 1.30 % | 87.5% |
| | **OPTIMAM** | 76.75 % ± 0.66 % | 76.9% |
| | **MIAS + DDSM** | 84.14 % ± 1.22 % | 84.79% |

method obtained a classification accuracy of $85.2\% \pm 5.7\%$ for a set of 300 cases. Later, Suhail et al. (2018b) developed a topological model using the binary tree properties and classified the clusters as benign/malignant based on the height of the tree and obtained an accuracy of 55.0% for the whole dataset and 91.0% for a subset of 129 images. To investigate the strength of topological modelling for micro-calcification classification, Ashiru and Zwiggelaar (2016) conducted a study to compare the topological modelling by Chen et al. (2015) and location-based classification by Andreadis et al. (2015). The study found that topological models performed better in explaining the micro-calcification clusters. Unlike other topological models discussed, we utilized the distributional arrangement of micro-calcifications in a cluster to form sub-chain-clusters to represent the relative arrangement of benign and malignant micro-calcifications. We obtained a classification accuracy of $86.47\% \pm 1.30\%$ for a set of 289 cases.

Though the proposed approach can classify most of the cases correctly, there were misclassified instances like the mammographic patches shown in Figure 5.12. The top row ROI was misclassified as a malignant case by the proposed approach, where the micro-calcifications were closely distributed. However, they were larger, and so they merged quickly to become a single cluster with minimum scaling which contradicts the assumption that benign calcifications are widely spread and needed a higher number of

Table 5.7 The area under the ROC curve ($A_z$) for the MIAS, DDSM and OPTIMAM datasets for 10 run 10-FCV and Leave-one-out approaches.

| | | $A_z$ | |
|---|---|---|---|
| | | **10-FCV** | **Leave-one-out** |
| **Dataset** | **MIAS** | $0.849 \pm 0.013$ | 0.833 |
| | **DDSM** | $0.893 \pm 0.008$ | 0.899 |
| | **OPTIMAM** | $0.779 \pm 0.005$ | 0.775 |
| | **MIAS + DDSM** | $0.880 \pm 0.010$ | 0.886 |

Table 5.8 Overview of our results with related works on same dataset.

| Feature | Database | Cases | Classifier | Method | Result |
|---|---|---|---|---|---|
| Intensity, shape and linear structures | DDSM | 150 | ANN/SVM | Ren (2012) | $A_z$=0.94 |
| Shape | DDSM | 183 | Threshold | Ma et al. (2010) | $A_z$=0.96 |
| Cluster | MIAS | 25 | SVM | Papadopoulos et al. (2005) | $A_z$=0.81 |
| Topology & location | MIAS | 20 | KNN | Ashiru and Zwiggelaar (2016) | $A_z$=0.95 |
| Topology & location | DDSM | 280 | KNN | Ashiru and Zwiggelaar (2016) | $A_z$=0.75 |
| Law features | MIAS | 322 | SVM | Dheeba and Selvi (2011) | CA=86.1 |
| Topological features | DDSM | 300 | KNN | Strange et al. (2014) | CA=80% |
| Tree-based modelling | DDSM | 129 | Majority voting | Suhail et al. (2018b) | CA=91% |
| Topology | DDSM | 300 | KNN | Chen et al. (2015) | CA=85.2 % ± 5.7 % |
| Connected-chain model | DDSM | 289 | KNN | George et al. (2018c) | CA=86% |
| Multi-scale connected-chain | DDSM | 289 | KNN | Ours | CA=86.47 % ± 1.30 %<br>$A_z$=0.893 ± 0.008 |
| Multi-scale connected-chain | MIAS | 20 | KNN | Ours | CA=82.5 % ± 2.63 %<br>$A_z$=0.849 ± 0.013 |
| Multi-scale connected-chain | OPTIMAM | 286 | KNN | Ours | CA=76.75 % ± 0.66 %<br>$A_z$=0.779 ± 0.005 |

scales before they overlap. This was unexpected for the proposed approach where we assumed that for benign cases, the micro-calcifications were widely spread. Similarly, the second row represents a malignant ROI reported as benign because the micro-calcifications were widely distributed though they were of small size. So with a limited number of scales, the merging of micro-calcifications was not possible, which led to the extraction of features similar to benign properties.

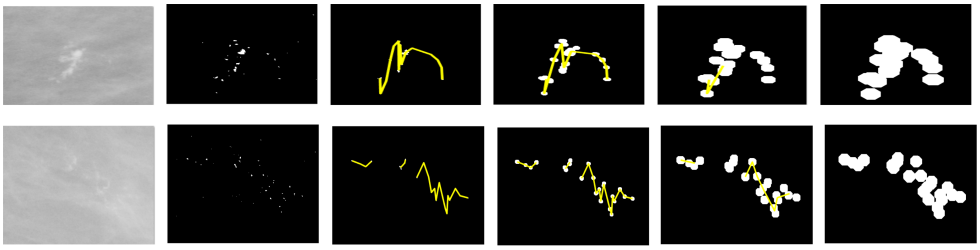


Fig. 5.12 Misclassified examples. First row: a benign ROI reported as malignant, second row: a malignant ROI reported as benign.

**Conclusion and Discussion**

The proposed novel approach of considering the topological distribution of micro-calcifications in mammogram ROIs using a multi-scale approach found to have comparable results to the literature. The method takes the clinical description of micro-calcification distribution into account for the cluster classification process (Henrot et al., 2014). The results obtained are similar with techniques reported in the literature for the MIAS and DDSM datasets, as discussed in the previous Section and Table 5.8. The proposed algorithm was implemented for a specific number of scales, and the optimum number of levels is an essential factor which will be investigated in the future. Similarly, additional datasets will be considered to evaluate the robustness of the algorithm. Simultaneously, the effect of the segmentation method on the classification results will be studied in more detail. In addition to the micro-calcification distribution features, the surrounding tissue characteristics and morphological traits can be extracted for additional discrimination between benign/malignant cases.

As most of the micro-calcification classification works in the literature uses statistical and morphological properties like individual micro-calcification features and cluster region features, the prominence of topological features over traditional approaches are to be investigated. Therefore, a investigative study on traditional texture, morphological and statistical features are compared with the topological features for micro-calcification classification in the next Section 5.5.

## 5.5 Combined Chain and Tissue Features for Micro-calcification Classification

In the previous sections, we have discussed in detail the proposed connected-chain modelling method for micro-calcification classification. There are various studies in the literature where the MC tissue features were investigated to classify micro-calcification's into benign or malignant. Here, the efficiency of topological distributional models and tissue features in classification of micro-calcifications are explored. Similarly, we investigate whether the shape of the tissue region will bring a large variation in the classification of micro-calcifications. So for this purpose, we extract convex and concave shapes surrounding tissue for texture analysis.

### 5.5.1 Data

The data used for the evaluation were from the DDSM as in the previous studies. We used DDSM as it has an appropriate size dataset and is used by other researchers,

which helps in providing a comparison with other methods. As discussed in previous studies, the distribution of malignant and benign MCCs differ. In case of malignant cluster, the MCs are closely packed with numerous number of MCs while in case of benign cluster, the MCs are widely distributed or has fewer number of MCs in a cluster. Since the concave and convex regions are analysed for the surrounding tissue region analysis, we choose only ROIs with a minimum of three MC's after the denoising procedure to form a region. The surrounding tissue regions are analysed along with the MC individual statistical and morphological features and connected-chain to find the effect of each feature on the classification of MCC into benign or malignant cases. After the selection of ROI's, the dataset contained 241 mammogram ROIs for feature extraction and classification (see to Section 2.9 for more details on the dataset).

### 5.5.2 Experimental Set-up

The experimental study evaluated the connectivity and distributional features of micro-calcifications in a segmented ROI and the texture of tissue around each micro-calcification cluster along with the statistical characteristics of individual MCs. From a clinical perspective, the distribution of micro-calcifications in benign cases are widely spread while the malignant micro-calcifications are closely located, forming a dense and compact cluster. We utilized these distributional characteristics of micro-calcification for constructing the connected-chain graph. Subsequently, the clusters formed through the connected-chain approach were extracted as objects to evaluate the tissue features around the micro-calcification clusters. Haralick features (Haralick et al., 1973) and M-ELBP features (see Chapter 3), were extracted for cluster texture analysis.

The input to the experimental study was histologically confirmed MCCs segmented ROIs. The input ROIs were pre-processed and binarized to perform binary morphological operations on the selected ROIs during processing. The distributional closeness of micro-calcifications was calculated through a connected-chain graph method forming chains of closely distributed micro-calcifications. Subsequently, to evaluate the effect of surrounding tissue, the texture features of MC cluster was computed. Similarly, the individual MC features like mean intensity, eccentricity, the distance between MCs, intensity features, elongation, the size and area of MCs were calculated. A correlation-based feature subset (CFS) selection method with a best-first-search method was used to select the unique features from the extracted feature vector. The detailed overview of the approach is shown in Figure 5.13.

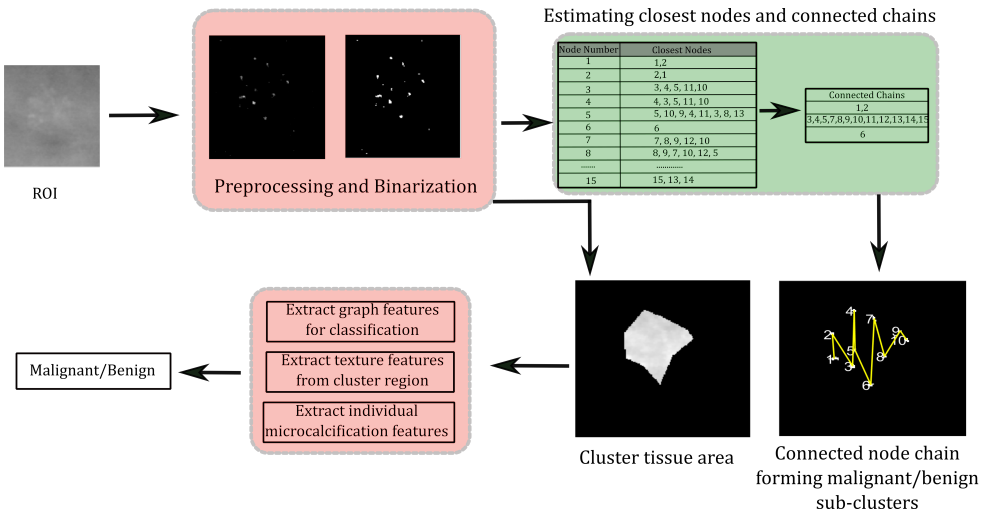



Fig. 5.13 Overview of feature extraction through connected chain and MC cluster surrounding tissue region.

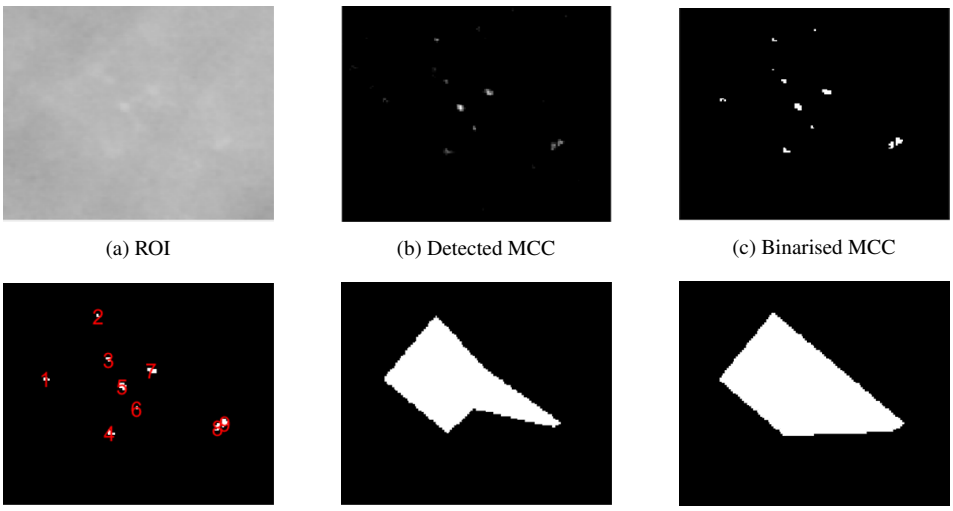
The connected chain model for investigating the distributional pattern are constructed as explained in Section 5.3.1.

### 5.5.3 Texture Feature Extraction from Surrounding Tissue

Various studies have shown that the texture properties of the tissue surrounding MCs can contribute to breast cancer diagnosis (Chan et al., 1998; Karahaliou et al., 2007). Generally speaking, the intensity distribution of MCCs is an important characteristic of micro-calcification segmentation. Therefore, some pattern recognition approaches use a grey-level co-occurrence matrix (GLCM) to extract characteristics features of patterns. The approach behind GLCM is to represent textures by a matrix of pair of grey-level appearing probabilities. The grey-level co-occurrence matrix (GLCM) was chosen to calculate the texture features of the cluster region formed by MCCs (Haralick et al., 1973; Karahaliou et al., 2007; Soltanian-Zadeh et al., 2001; Thiele et al., 1996). We used four angles (0°, 45°, 90°, 135°). The Haralick texture features extracted at all four angles included energy, contrast, correlation, homogeneity, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, maximal correlation coefficient and two information measurement of correlations.

#### Surrounding Tissue Region Extraction

In-order to evaluate the effect of tissue region for classification, we evaluated the concave and convex tissue regions around the MCC. The difference in shape of the tissue regions are shown in Figure 5.14.



(a) ROI (b) Detected MCC (c) Binarised MCC

(d) Labelled MC centroids (e) Concave region (f) Convex region

Fig. 5.14 Extraction of concave and convex regions of MC cluster for tissue texture feature evaluation.

We performed the cluster morphological feature extraction from both the concave hull region and convex region. We found that there is not much variation in classification accuracy by the shape of the region as both regions could include healthy tissues for feature extraction. The features extracted included area, eccentricity, orientation, perimeter, major-axis and minor-axis length of the regions along with few intensity features of the region like max, mean, minimum intensity of the region and number of MC in the region. The feature selection was performed, and the features like major-axis length, intensity features and number of MCs were selected as a prominent feature for classification. The Bayesian Network classifier was used to estimate the performance of the chosen features. The concave cluster features obtained a CA% of 73.31% with an $A_z$ of 0.756 while the convex cluster features obtained a classification accuracy of $CA = 73.72\%$ with an $A_z$ of 0.751. The results showed that there is not much variation in CA% and $A_z$ with the shape features of the cluster region. Still, the number of MCs, intensity, length features have more potential effect in classification.

Therefore for further experimental studies, we chose the convex region for texture feature extraction and classification. For texture feature extraction, tissue around the cluster is segmented as shown in Figure 5.15

### 5.5.4 Individual Micro-Calcification Features

In our experimental study, we extract morphological and gray level features. The set of features included the features like minimum and mean distance between MCs, total

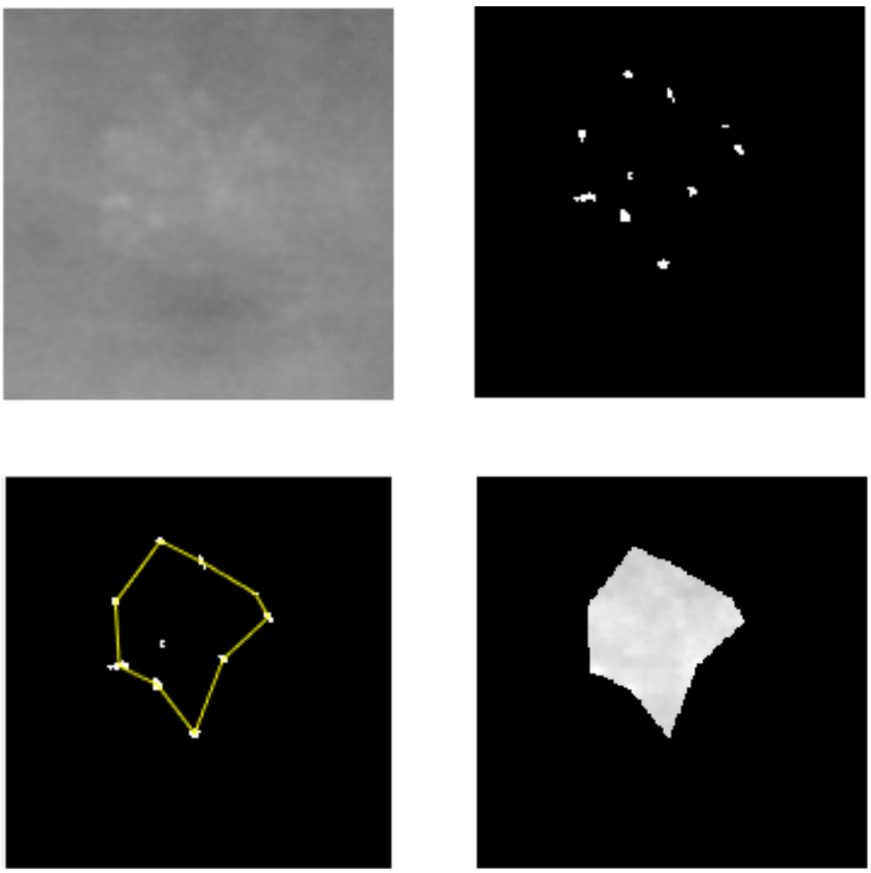


Fig. 5.15 Extraction of MC cluster region for tissue texture feature evaluation.

area covered by MCs, eccentricity, elongation, the mean intensity level and standard deviation of the intensity level were selected after feature selection.

The comparative study between the connected-chain features, cluster tissue region and individual micro-calcifications were performed alone and in combination. To estimate the potential of individual feature set and combination of feature sets, the selected features were classified using the Bayesian Network and AdaBoost classifiers.

### 5.5.5 Experimental Results and Discussion

To explore the texture, topological distributional features and individual MC features, we use histologically proved clusters ROIs segmented from the DDSM database. The chain-connected sub-clusters were evaluated using AdaBoost and Bayesian classifiers using a ten-run ten-fold cross-validation for chain features. The best classification accuracy obtained using Adaboost was 84.23% with an $A_z$= 0.87 with a sensitivity of 0.85, a specificity of 0.84 and F-measure of 0.84. While the 10 run-10 fold cross-validation obtained a CA of $83.734\% \pm 0.381$ and $A_z = 0.864 \pm 0.007$. With

the Bayesian classifier, the connected-chain features obtained a CA of 82.99% with $A_z$=0.86, specificity and sensitivity of 0.83. While the 10 run-10 fold cross-validation obtained a CA of $82.448\% \pm 0.519$ and $A_z = 0.854 \pm 0.007$ as shown in Table 5.9.

While considering the morphological and intensity features for individual MC features, the features selected were minimum and mean distance between MC, the area covered by MCs, the standard deviation of eccentricity, elongation features, the intensity features like mean, min and max intensity of MCs. Using the selected features, the classification done by AdaBoost obtained the best classification of CA=74.57%, $A_z$=0.79, sensitivity of 0.76, specificity, and F-measure to be 0.75. The 10-run 10-FCV obtained a CA=$72.425\% \pm 1.102\%$ and $A_z = 0.778 \pm 0.010$. Similarly, with the Bayesian Network, it achieved the best CA of 75.0%, $A_z = 0.82$ with a specificity of 0.74, sensitivity of 0.75, and F-measure to be 0.75. The 10-run 10-FCV achieved CA=$74.237\% \pm 1.034$ and $A_z = 0.813 \pm 0.006$.

While considering the surrounding tissue of the MC cluster, we extracted the Haralick features of the concave pattern of the cluster, as shown in Figure 5.15. With the AdaBoost classifier, the selected Haralick features obtained the best CA=75.52% with $A_z$=0.78, a sensitivity of 0.75, specificity and F-measure to be 0.76. The ten-run ten-FCV achieved a CA of $75.269\% \pm 0.981\%$ and $A_z = 0.777 \pm 0.009$. Similarly, with the Bayesian Network, the best CA achieved was 75.52% with $A_z$ =0.77, specificity and sensitivity of 0.76 and F-measure of 0.75. The average CA achieved for a 10 run 10 FCV was $72.904\% \pm 1.341\%$ and $A_z = 0.764 \pm 0.010$. It was noted that the strongest features among the Haralick features in classification were contrast, variance and maximal correlation coefficient. In addition to Haralick features, we explored the M-ELBP features for the convex regions attaining a CA% to be $67.887\% \pm 0.612\%$ and $A_z = 0.684 \pm 0.013$ with specificity of 0.68, sensitivity and F-measure to be 0.69. The poor results of M-ELBP could be due to the smaller tissue convex region, which results in a lack of texture information for the operator. Therefore, we focus more on Haralick feature for the combination study as it is used as the standard feature extraction by other researchers.

Each feature vector was combined to investigate the efficiency of classification. A 10-run 10-FCV was performed on the combined features and found that the classification accuracy and ROC increase with the feature combination. A detailed table for the classification accuracy and $A_z$ can be seen in Table 5.9

The classification results in Table 5.9 shows a promising result with the combination of a distributional model with texture and individual MC features. To statistically analyse the classification on the feature combination by classifiers, each selection or the 10-FCV was repeated ten times to obtain 100 classifier models (10 folds$\times$ ten runs).

Table 5.9 The classification accuracy and $A_z$ for the Bayesian Network and AdaBoost classifiers.

| | CA% | | $A_z$ | |
|---|---|---|---|---|
| | **Bayesian Network** | **AdaBoost** | **Bayesian Network** | **AdaBoost** |
| **Connected-chain** | $82.448\% \pm 0.519\%$ | $\mathbf{83.734\% \pm 0.381\%}$ | $0.854 \pm 0.007$ | $0.864 \pm 0.007$ |
| **M-ELBP texture** | $67.887\% \pm 0.612\%$ | $68.457\% \pm 0.861\%$ | $0.684 \pm 0.013$ | $0.700 \pm 0.012$ |
| **Individual MC features** | $74.237\% \pm 1.034\%$ | $72.425\% \pm 1.102\%$ | $0.813 \pm 0.006$ | $0.778 \pm 0.010$ |
| **Cluster texture /GLCM** | $72.904\% \pm 1.341\%$ | $75.269\% \pm 0.981\%$ | $0.764 \pm 0.010$ | $0.777 \pm 0.009$ |
| **GLCM and Individual MC** | $79.576\% \pm 1.227\%$ | $79.364\% \pm 0.663\%$ | $0.863 \pm 0.852$ | $0.815 \pm 0.006$ |
| **GLCM and Connected chain** | $77.800\% \pm 1.177\%$ | $81.535\% \pm 1.361\%$ | $0.864 \pm 0.006$ | $0.893 \pm 0.008$ |
| **Individual MC and Connected chain** | $82.167\% \pm 0.479\%$ | $81.701\% \pm 1.676\%$ | $0.911 \pm 0.004$ | $0.889 \pm 0.011$ |
| **GLCM, Individual MC and Connected chain** | $81.991\% \pm 0.710\%$ | $80.705\% \pm 1.977\%$ | $\mathbf{0.912 \pm 0.003}$ | $0.815 \pm 0.006$ |

The classification accuracy and the area under the ROC curve $A_z$ was calculated for all combinations. It was found that the topological distributional methods perform better than all other combinations with a CA of $83.734\% \pm 0.381\%$ and $A_z = 0.864 \pm 0.007$ by the AdaBoost classifier with average sensitivity and specificity of 0.84.

Similarly, the classification of feature combinations with chain features performed better than with tissue texture or individual MC features. The area under the ROC curve $A_z$ has improved from $0.854 \pm 0.007$ and $0.813 \pm 0.006$ for chain feature and individual MC feature respectively to $0.911 \pm 0.004$ by the Bayesian Network. Similarly, for AdaBoost classifier, the $A_z$ improved from $0.864 \pm 0.007$ and $0.778 \pm 0.010$ to $0.889 \pm 0.011$. Similarly, the $A_z$ of individual MC, texture feature, the connected-chain had improved from $0.854 \pm 0.007$, $0.813 \pm 0.006$, $0.764 \pm 0.010$ to $0.912 \pm 0.003$ when the features got combined and classified using the Bayesian Network classifier. The experimental study shows the importance of investigating the MC distributional structure pattern in a cluster using the surrounding tissue.

To evaluate the performance of feature combination in detail, the classification accuracy by both classifiers are plotted in Figure 5.16 and Figure 5.17.

The graph shows that the connected-chain features alone have the potential in discriminating the malignant and benign MC cluster with both classifiers. Similarly, the surrounding tissue and individual micro-calcification features tend to have similar classification behaviour when comparing the results of the classifiers.

Classification accuracy is not the single factor that determines the potential of features or classifiers in a CAD system. Therefore the area under the curve, $A_z$ is calculated for each run for the ten-run 10-FCV for the Bayesian network classifier and Adaboost as shown in Figure 5.18 and Figure 5.19.

It is to be noted that the $A_z$ value has been improved when the distributional pattern features are combined with tissue texture features and the individual MC features for the Bayesian network. The texture features of the surrounding tissue have the least $A_z$ value as it could be because of the appearance of healthy tissue in the cluster region.

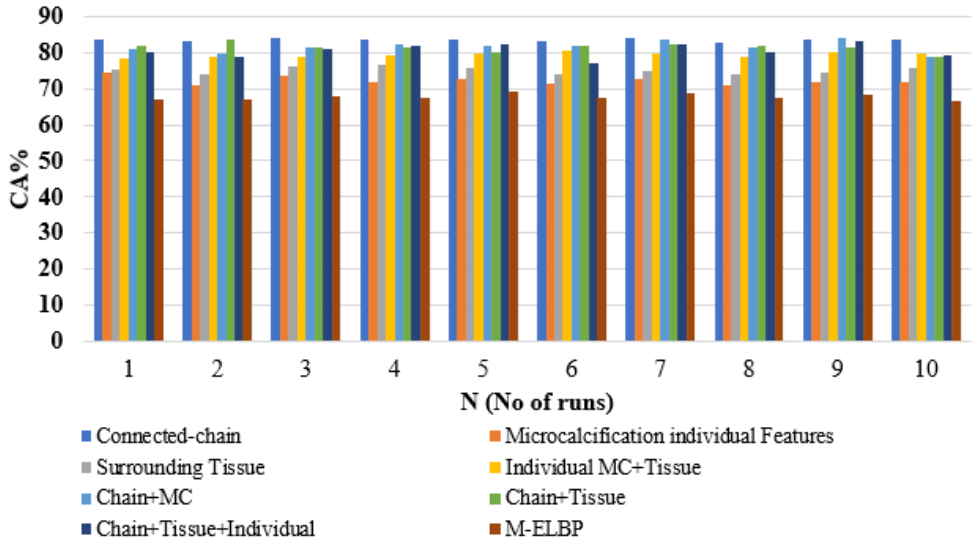



Fig. 5.16 Classification accuracy (CA%) of feature combinations with respect to each run in 10 run 10 FCV using AdaBoost classifier

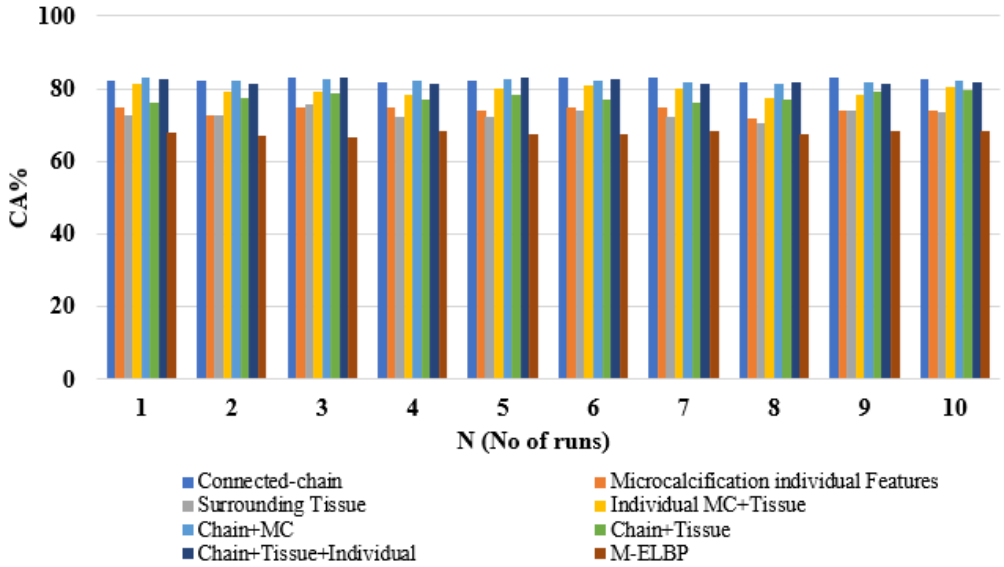



Fig. 5.17 Classification accuracy (CA%) of feature combinations with respect to each run in 10 run 10 FCV using the Bayesian Network classifier

The combination of chain features with individual MC statistical features and all the three feature set combinations obtained more than 0.90 in every run, improving the average $A_z$ to $0.911 \pm 0.004$ and $0.912 \pm 0.003$ respectively.

The $A_z$ value by AdaBoost classifier has increased above 0.90 when all the features are combined with an average of $0.893 \pm 0.008$ for a feature combination of tissue and

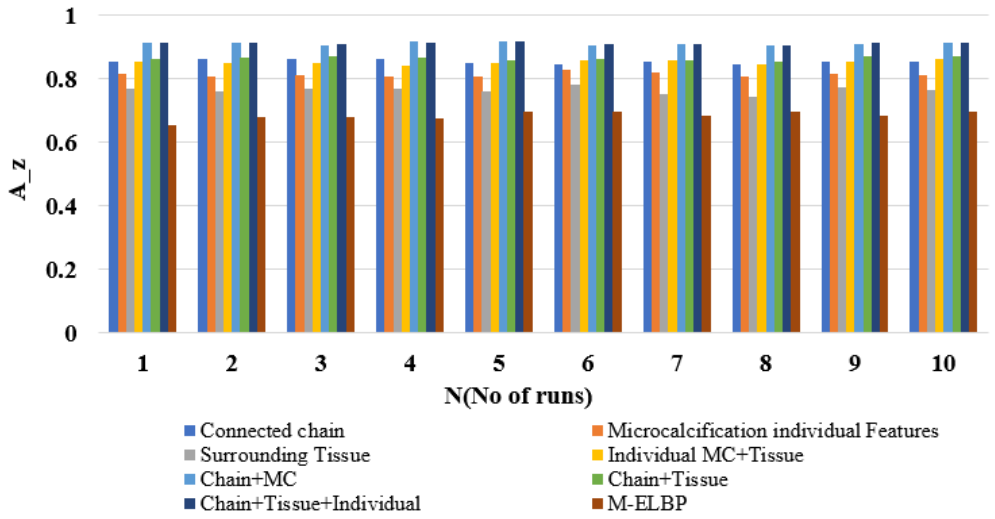



Fig. 5.18 The $A_z$ for feature combinations with respect to each run in 10-run 10-FCV using the Bayesian Network classifier.

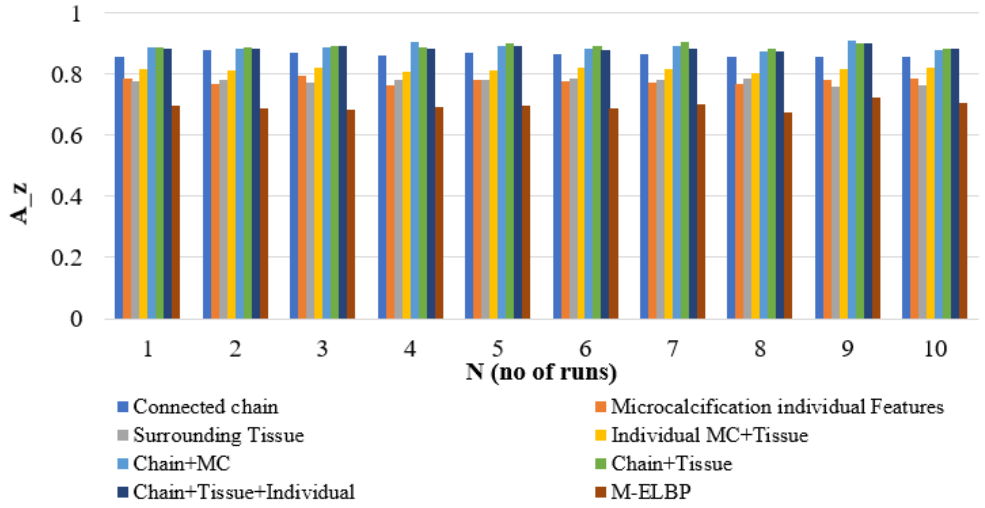



Fig. 5.19 The $A_z$ for feature combinations with respect to each run in 10 run 10 FCV using the AdaBoost classifier.

individual micro-calcifications and $0.889 \pm 0.011$ for chain and individual statistical features. While comparing the classification accuracy and $A_z$ values, we can predict that the tissue or statistical features along the distributional features could improve the feature discrimination power of the CAD system.

To investigate the results further, we have compared similar state of the art studies. Table 5.10, represent the various studies related to micro-calcification classification studies combining different approaches. The study found that connected chain topo-

logical model performed better than the texture or individual MC feature set. It is noted that with the addition of more features, for instance, the CA% decreases slightly as in Figure 5.16 and Figure 5.17. The studies show that for a fixed number of data points/objects, the accuracy increases then decreases for increasing numbers of features (Hua et al. (2005); Hughes (1968); Jensen and Shen (2009); Trunk (1979)).

Table 5.10 Qualitative comparison of our results with those achieved by related work.

| Feature | Dataset | Cases | Classifier | Method | Result |
|---|---|---|---|---|---|
| Texture+cluster | Unknown | 191 | ANN | Dhawan et al. (1996) | $A_z$=0.86 |
| Shape+cluster | Liverpool | 38 | KNN | Betal et al. (1997) | $A_z$=0.79,$A_z$=0.84 |
| Cluster+morphology | Univ of Chicago | 104 | SVM | Wei et al. (2009) | $A_z$=0.82 |
| Topology+location | MIAS<br>DDSM | 20<br>280 | KNN | Ashiru and Zwiggelaar (2016) | $A_z$=0.95(MIAS),<br>$A_z$=0.75(DDSM) |
| Tree-based modelling | DDSM | 129 | Voting | Suhail et al. (2018b) | CA=91% |
| Topology | DDSM | 80 | KNN | Chen et al. (2015) | CA=95%,$A_z$=0.96 |
| Connected-chain+texture | DDSM | 241 | AdaBoost | Ours | $A_z = 0.893 \pm 0.008$ |
| Connected-chain+Individual MC | DDSM | 241 | Bayesian Network | Ours | $A_z = 0.911 \pm 0.004$ |
| Connected-chain+Individual MC+Texture | DDSM | 241 | Bayesian Network | Ours | $A_z = 0.912 \pm 0.003$ |

Table 5.10, indicates that the results obtained through the proposed approach was comparable with alternative approaches. The presented study along with Ashiru and Zwiggelaar (2016) shows that the connectivity/topological distribution is an important research area to be explored further for micro-calcification classification.

## 5.6 Summary

Micro-calcifications are small deposits of calcium seen as white specks in mammographic images. Detection and classification of benign and malignant microcalcifications is an essential issue for CAD systems as it can assist in the early diagnosis and treatment of breast cancer. CAD systems with high accuracy for detection and classification can act as a second reader of mammographic images and can reduce incorrect treatment. A novel MCC classification method called connected-chain model based on the distributional patterns of MCs in the cluster was proposed and found comparable result with state of the art methods. Similarly, the multi-scale connected-chain model was introduced to analyse the effect of distribution and resolution on MC classification using a connected-chain method.

Similarly, a comparative study to estimate the potential of the proposed method with tissue texture and morphology of the MCs was done. The study showed that a combination of tissue features or MC features along with the distributional pattern

analysis features could improve the potential of a CAD system in discriminating the micro-calcifications as benign or malignant.

# Chapter 6

# Conclusion and Future Work

In this final chapter, we present the summary of this thesis and the conclusions drawn from the work. Moreover, we summarise the main contributions and applications of the work and novel aspects of our research in the field of mammographic image analysis. Towards the end of this chapter, a list of our publications related to this thesis is provided.

## 6.1 Summary and Conclusion of the Thesis

In this thesis, we have developed a set of mammographic image analysis methods using computer vision and image processing techniques. The image analysis methods explored various mammographic image analysis approaches for parenchymal patterns modelling, risk assessment, and breast density classification. Besides, an automated micro-calcification classification method has been proposed to classify the segmented MCC as benign or malignant clusters.

The thesis started (Chapter 1) with a study on the current breast cancer statistics worldwide showing the necessity and possibility of more research in breast cancer diagnosis. Moreover, brief descriptions on different breast imaging modalities, breast cancer risk factors, various breast density classification schemes have been covered. Furthermore, a detailed study on the nature of micro-calcification and its clinical features used for discriminating the benign and malignant cases were covered. Following these, we set the basis for our thesis work on risk assessment and parenchymal pattern modelling, breast density classification and micro-calcification discrimination as benign or malignant. Following this (Chapter 2), we carried out a literature review on the role of CAD systems in medical image analysis and studied existing approaches in the literature. We performed a detailed study on various existing methods for breast

density scoring and estimation, followed by micro-calcification descriptors and its different classification approaches used in the literature. From all these studies in the literature, we concluded that it is still worth developing methods for mammographic image analysis on risk estimation, breast density estimation and MCC classification although there were many methods developed in the past decade. The studies established a basis for our work in the thesis. So, for the image analysis methods and feature extraction approaches related to the works presented in the thesis, a detailed study on texture, its uses and various descriptors for feature extraction were explored. Moreover, a detailed description of the datasets used in the thesis was presented towards the end of Chapter 2.

For mammographic density estimation (Chapter 3), we developed a topology-based LBP variant named Mean-Elliptical Local Binary Pattern (M-ELBP) for BI-RADS based and tissue-based breast density classification. Unlike traditional LBP or ELBP, the method explored the mean intensity value of neighbourhood pixels along with texture information to generate the binary pattern. The classification accuracy performance of the proposed method was tested on two datasets, DDSM and MIAS using tissue-based (77.38% for MIAS) and BI-RADS (75.4% for MIAS and 47.65% for DDSM) based ground truth information. The proposed method was tested against various traditional texture analysis methods like GLCM, LBP, ELBP, U-ELBP, and LDP for performance evaluation. Moreover, a study on the descriptor size effect, ROI size effect, and the role of classifiers in density estimation was explored. Based on our experimental investigations, it was found that the proposed method performed better compared to traditional methods and was more robust for breast density estimation.

For breast parenchymal modelling and risk estimation (Chapter 4), a new multi-scale elliptical blob-based approach was developed. We focused on approximately blob-like tissue patterns in mammograms and used elliptical multi-scale blobs to represent the tissue patterns. Distance-based and qualitative relations among blobs were used to merge the overlapping blobs. The performance of the approach was tested on the MIAS dataset for discriminating between high/low risk (CA% = 66.13). Besides, multi-scale elliptical blobs were used to represent different orientations of tissue patterns in mammograms.

For mammographic MCC classification (Chapter 5), a new method called connected-chain was proposed for discriminating segmented MCC as benign or malignant MCC. The connectivity/closeness of micro-calcifications in MCC was explored in the proposed approach. The features were extracted from a single scale and also from multi-scale connected-chains to classify MCCs as benign or malignant. The validity of the method was tested using three datasets, MIAS, DDSM and OPTIMAM. The

method found to be promising and comparable with other topological methods in the literature using the same dataset (86.47%, 82.50%, and 76.75% for DDSM, MIAS and OPTIMAM, respectively ).

In summary, we conclude that the methods proposed and the results obtained in the thesis have the potential for classification of breast density and MCCs. While considering the multi-scale elliptical blob representation of tissue patterns, the visual evaluation of blobs at different orientation and scale showed the possibility of representing and evaluating the nature of parenchymal patterns for risk estimation. Similarly, using intensity features along with texture features provided a better classification for the M-ELBP operator. Both the multi-scale elliptical blob study and the M-ELBP method showed that intensity information plays an essential role in breast tissue density characterisation/classification. The study showed that exploring more intensity related features can improve the performance and efficiency when combined with spatial (texture) information in discriminating the tissue type. Moreover, the connected-chain approach for MCC classification showed the possibility of exploring more topological/distributional features in determining the nature of MCC. The method found to be promising and comparable with related topological works in the literature. In short, the methods proposed in this thesis found to be efficient in representation, discrimination of breast tissue density and MCC classification.

### 6.1.1 Alternative Application Area

The main objective of the thesis was to analyze mammographic images and develop a model for parenchymal pattern representation, breast density estimation and MCC classification. From the research studies and experimental set-ups, we proposed a multi-scale elliptical blob model for representing the parenchymal breast patterns. The proposed method was successful to an extent in representing the orientation of parenchymal tissue in a mammogram, although it needs further development to be used as a clinical application. The method can be translated in representing different size objects in an image at different orientations and size like in historical photos, satellite images. Similarly, the process can be utilized in other computer vision applications such as nuclei detection, object tracking, precise particle detection, blood vessel reconstruction, and phantom generation.

We proposed a texture feature extraction descriptor named M-ELBP as a modification to ELBP by incorporating the mean intensity value of the neighbouring pixel to generate a binary pattern. The method was found to be successful in breast density classification using MIAS and DDSM dataset. Unlike traditional LBP, where the

descriptor fails with noise, the proposed descriptor found to perform better with noise as there was no direct comparison between pixels. Moreover, the addition of the mean intensity features with texture features enhanced the feature set for better classification. Other than breast density classification, the proposed method can find various other application like face expression recognition, in textile industries for determining the type and quality of rugs, identifying and classifying regions from satellite images, and crowd density estimation.

For the classification of MCCs, we proposed a connected-chain method where the micro-calcifications that were closely packed/distributed were connected to each other as a chain, based on a threshold distance. The method found to be successful in its multi-scale/level approach in discriminating MCCs as benign or malignant. The method explored the distributional/topological pattern of micro-calcifications in the MCC. Other than the mentioned application, the proposed approach can be utilized in other image processing applications like building a cluster region and isolating the region in case of pandemic disease by evaluating the distance between positive cases, in generating a topological pattern of a region, and in generating the route map of moving objects or tracking moving objects through specific chain patterns.

## 6.2 Future Work

In this section, we outline a few aspects which could be investigated further in future research based on the work presented in this thesis:

- While developing the breast density segmentation methods, instead of considering the image density class, the segmented regions could be evaluated based on radiologist-annotated images.
- While considering the risk assessment, linear structures and other similar geometrical patterns could be regarded as features for assessment in addition to the texture or parenchymal region density.
- Considering the local feature extraction for density estimation, M-ELBP and LDP indicated better result which indicated the potential of directional filter methods. So, more directional texture features and gradient feature extractors could be considered for breast density estimation. Similarly, the addition of more intensity features into existing descriptors could be studied.
- Though the studies showed that the fibroglandular disc region provides better density classification results than the whole mammograms, extracting the ROI

could be investigated in a way to include maximum fibro-glandular areas for classification.

- While considering the connected chain method, an automated approach to detect the appropriate threshold for distance estimation based on dataset resolution can be investigated to make the process independent of dataset resolution parameters. Similarly, the effect of the segmentation method on classification result could be explored.
- While considering the blob analysis method, we used a Laplacian of Gaussian detector which gives strong responses at the edges and boundaries. Instead of LOG operator, other operators like Hessian-Laplacian or related operators could be used as they will not produce strong extrema in boundaries and can detect larger blobs than LOG making it more computationally cost-effective.
- While merging the elliptical blobs at different orientations, we used area and distance-based methods for removing overlapping blobs. A better merging could improve the classification because, during the image processing steps, certain blobs with a strong response get deleted due to blob size sorting in the merging process. Instead, the intensity features inside the blobs at different orientations at a particular region could be investigated and used for merging.
- While calculating the blob direction, automatic scale selection, and orientation selection for the blob from the extrema could be investigated to avoid the problem of blob merging based on scale and orientation.
- In this thesis, some of the works are based on digitized datasets. So, the studies could be extended on large digital datasets to estimate the robustness of the method.

## 6.3 Novel Contributions

- A new method for breast density estimation was done using the proposed novel method, mean-elliptical local binary pattern (M-ELBP). The proposed method incorporated intensity features into texture features to improve the feature extraction property of the operator.
- The study investigated the variation in classification results for breast density classification for the whole breast region and the ROI extracted from the fibro-glandular disc region.

- In addition to the selection of ROI from the fibro-glandular region, the size for ROI selection for the tested dataset was studied.
- To study the effect of the classifier on the classification process for the same set of features, we compared many classifiers with the same feature set and same feature selection method.
- The effect of descriptor size for feature extraction in the classification of breast density was investigated for the M-ELBP operator.
- For categorizing micro-calcification into benign and malignant, we introduced a new approach based on the topological distributional connectivity of MCs in MCC. The method was called connected-chain model where the closely distributed MCs are linked to form a connected-chain. The chain properties were later used for classification.
- To reduce the effect of the resolution of mammograms from different dataset to an extent, we performed a multi-scale model of connected-chain for MC classification.
- To investigate the efficiency of the proposed connected-chain approach, we compared the method with existing feature extraction factors like surrounding tissue texture analysis of MCs and individual MC morphological and statistical features.
- For investigating the effect of MCC surrounding tissue structure shape, we extracted convex and concave shapes for MC classification.
- For the mammographic risk estimation, we modelled the mammographic parenchymal structures using elliptical blobs in different orientations.

## 6.4 List of Publications

- **Journal Publications**
- Qu, Y., Fu, Q., Shang, C., Deng, A., Zwiggelaar, R., George, M. and Shen, Q., 2020. Fuzzy-rough assisted refinement of image processing procedure for mammographic risk assessment. Applied Soft Computing, p.106230.

- George, M. and Zwiggelaar, R., 2019. Comparative study on local binary patterns for mammographic density and risk scoring. Journal of Imaging, 5(2), p.24.

- George, M., Chen, Z. and Zwiggelaar, R., 2019. Multiscale connected chain topological modelling for microcalcification classification. Computers in Biology and Medicine, 114, p.103422.

- **Conference publications**

- George, M., Denton, E.R. and Zwiggelaar, R., 2018, September. Topological Connected Chain Modelling for Classification of Mammographic Microcalcification. In CGVC (pp. 1-5).

- George, M., Denton, E. and Zwiggelaar, R., 2018, July. Mammogram breast density classification using mean-elliptical local binary patterns. In 14th International Workshop on Breast Imaging (IWBI 2018) (Vol. 10718, p. 107180B). International Society for Optics and Photonics.

- George, M. and Zwiggelaar, R., 2018, July. Breast tissue classification using Local Binary Pattern variants: A comparative study. In Annual Conference on Medical Image Understanding and Analysis (pp. 143-152). Springer.

- George, M., Denton, E. and Zwiggelaar, R., 2016. Mammographic Ellispe Modelling for Risk Estimation. Procedia Computer Science, 90, pp.163-168.

- George, M., Rampun, A., Denton, E. and Zwiggelaar, R., 2016, June. Mammographic ellipse modelling towards birads density classification. In International Workshop on Breast Imaging (pp. 423-430). Springer.

# References


Solmaz Abbasi and Farshad Tajeripour. Detection of brain tumor in 3d mri images using local binary patterns and histogram orientation gradient. *Neurocomputing*, 219:526–535, 2017.

ACR. Acr breast imaging reporting and data system. *Breast Imaging Atlas, 4th edn., Reston, VA, USA*, 2003a.

ACR. Breast imaging reporting and data system. *BI-RADS*, 2003b.

DD Adler and MA Helvie. Mammographic biopsy recommendations. *Current Opinion in Radiology*, 4(5):123–129, 1992.

Chul Kyun Ahn, Changyong Heo, Heongmin Jin, and Jong Hyo Kim. A novel deep learning-based approach to high accuracy breast density estimation in digital mammography. In *Medical Imaging 2017: Computer-Aided Diagnosis*, volume 10134, page 101342O. International Society for Optics and Photonics, 2017.

Narendra Ahuja. Dot pattern processing using voronoi neighborhoods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (3):336–343, 1982.

Raid Rafi Omar Al-Nima, Mohammed AM Abdullah, Musab TS Al-Kaltakchi, Satnam Singh Dlay, Wai Lok Woo, and Jonathon A Chambers. Finger texture biometric verification exploiting multi-scale sobel angles local binary pattern features and score-based fusion. *Digital Signal Processing*, 70:178–189, 2017.

Nashid Alam, Arnau Oliver, Erika RE Denton, and Reyer Zwiggelaar. Automatic segmentation of microcalcification clusters. In *Annual Conference on Medical Image Understanding and Analysis*, pages 251–261. Springer, 2018.

Nashid Alam, Erika RE Denton, and Reyer Zwiggelaar. Classification of microcalcification clusters in digital mammograms using a stack generalization based classifier. *Journal of Imaging*, 5(9):76, 2019.

Edén A Alanís-Reyes, José L Hernández-Cruz, Jesús S Cepeda, Camila Castro, Hugo Terashima-Marín, and Santiago E Conant-Pablos. Analysis of machine learning techniques applied to the classification of masses and microcalcification clusters in breast cancer computer-aided detection. *Journal of Cancer Therapy*, 3(06):1020, 2012.

Marwah Thamer Ali, Faiez Musa Lahmood, and Raad Farhood Chisab. Mammogram image segmentation for improving the diagnosis of dense breast issues. *International Journal of Engineering & Technology*, 8(1):44–52, 2019.

Douglas G Altman and J Martin Bland. Diagnostic tests. 1: Sensitivity and specificity. *BMJ: British Medical Journal*, 308(6943):1552, 1994.

Ioannis I Andreadis, George M Spyrou, and Konstantina S Nikita. A cadx scheme for mammography empowered with topological information from clustered microcalcifications' atlases. *IEEE Journal of Biomedical and Health Informatics*, 19(1): 166–173, 2015.

Adriana Antonelli, Gustavo Meschino, and Virginia Ballarin. Mammographic density estimation through permutation entropy. In *World Congress on Medical Physics and Biomedical Engineering 2018*, pages 135–141. Springer, 2019.

Sundaram Arivazhagan, L Ganesan, and S Padam Priyal. Texture classification using gabor wavelets based rotation invariant features. *Pattern Recognition Letters*, 27 (16):1976–1982, 2006.

Oluwaseun Ashiru and Reyer Zwiggelaar. Classification of mammographic microcalcification clusters using a combination of topological and location modelling. In *Image Processing Theory Tools and Applications (IPTA)*, pages 1–6. IEEE, 2016.

MR Ataollahi, J Sharifi, MR Paknahad, and A Paknahad. Breast cancer and associated factors: a review. *Journal of Medicine and Life*, 8(Spec Iss 4):6, 2015.

Philippe Autier and Mathieu Boniol. Mammography screening: A major issue in medicine. *European Journal of Cancer*, 90:34–62, 2018.

Robert Azencott, Jia-Ping Wang, and Laurent Younes. Texture classification using windowed fourier filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):148–153, 1997.

Morteza Babaie, Shivam Kalra, Aditya Sriram, Christopher Mitcheltree, Shujin Zhu, Amin Khatami, Shahryar Rahnamayan, and Hamid R Tizhoosh. Classification and retrieval of digital pathology scans: a new dataset. In *CVMI workshop of CVPR*, 2017.

Jean Babaud, Andrew P Witkin, Michel Baudin, and Richard O Duda. Uniqueness of the gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (1):26–33, 1986.

Min Sun Bae, Hyeong-Gon Moon, Wonshik Han, Dong-Young Noh, Han Suk Ryu, In-Ae Park, Jung Min Chang, Nariya Cho, and Woo Kyung Moon. Early stage triple-negative breast cancer: imaging and clinical-pathologic factors associated with recurrence. *Radiology*, 278(2):356–364, 2015.

Jay A Baker, Eric L Rosen, Joseph Y Lo, Edgardo I Gimenez, Ruth Walsh, and Mary Scott Soo. Computer-aided detection (cad) in screening mammography: sensitivity of commercial cad systems for detecting architectural distortion. *American Journal of Roentgenology*, 181(4):1083–1088, 2003.

Jay A Baker, Joseph Y Lo, David M Delong, and Carey E Floyd. Computer-aided detection in screening mammography: variability in cues. *Radiology*, 233(2): 411–417, 2004.

Predrag R Bakic, Michael Albert, Dragana Brzakovic, and Andrew DA Maidment. Mammogram synthesis using a 3d simulation. ii. evaluation of synthetic mammogram texture. *Medical Physics*, 29(9):2140–2151, 2002a.

Predrag R Bakic, Michael Albert, Dragana Brzakovic, and Andrew DA Maidment. Mammogram synthesis using a 3d simulation. ii. evaluation of synthetic mammogram texture. *Medical Physics*, 29(9):2140–2151, 2002b.

Isaac N Bankman, John Tsai, Dong W Kim, Olga B Gatewood, and William R Brody. Detection of microcalcification clusters using neural networks. In *Proceedings of 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 1, pages 590–591. IEEE, 1994.

LW Bassett, IA Hirbawi, N DeBruhl, and MK Hayes. Mammographic positioning: evaluation from the view box. *Radiology*, 188(3):803–806, 1993.

Craig A Beam, Peter M Layde, and Daniel C Sullivan. Variability in the interpretation of screening mammograms by us radiologists: findings from a national sample. *Archives of Internal Medicine*, 156(2):209–213, 1996.

Rhen Anjerome Bedruz, Edwin Sybingco, Argel Bandala, Ana Riza Quiros, Aaron Christian Uy, and Elmer Dadios. Real-time vehicle detection and tracking using a mean-shift based blob analysis and tracking approach. In *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, pages 1–5. IEEE, 2017.

Azeddine Beghdadi and Alain Le Negrate. Contrast enhancement technique based on local detection of edges. *Computer Vision, Graphics, and Image Processing*, 46(2): 162–174, 1989.

Alan Joseph Bekker, Hayit Greenspan, and Jacob Goldberger. A multi-view deep learning architecture for classification of breast microcalcifications. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*, pages 726–730. IEEE, 2016.

Peter N Belhumeur, João P Hespanha, and David J Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (7):711–720, 1997.

Rami Ben-Ari, Aviad Zlotnick, and Sharbell Hashoul. A weakly labeled approach for breast tissue segmentation and breast density estimation in digital mammography. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 722–725. IEEE, 2016.

H Berment, V Becette, M Mohallem, F Ferreira, and P Chérel. Masses in mammography: What are the underlying anatomopathological lesions? *Diagnostic and Interventional Imaging*, 95(2):124–133, 2014.

Donald A Berry, Giovanni Parmigiani, Juana Sanchez, Joellen Schildkraut, and Eric Winer. Probability of carrying a mutation of breast-ovarian cancer gene brca1 based on family history. *Journal of the National Cancer Institute*, 89(3):227–237, 1997.

Dibendu Betal, N Roberts, and GH Whitehouse. Segmentation and numerical analysis of microcalcifications on mammograms using mathematical morphology. *The British Journal of Radiology*, 70(837):903–917, 1997.

JC Bezdek. Pattern recognition with fuzzy objective algorithms plenum press. *New York*, pages 309–321, 1981.

Tushar Bhangale, Uday B Desai, and Upendra Sharma. An unsupervised scheme for detection of microcalcifications on mammograms. In *Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101)*, volume 1, pages 184–187. IEEE, 2000.

Christopher M Bishop. *Neural networks for pattern recognition*. Oxford university press, 1995.

Dorothea Blostein and Narendra Ahuja. Shape from texture: Integrating texture-element extraction and surface estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(12):1233–1251, 1989.

Lilian Blot and Reyer Zwiggelaar. Background texture extraction for the classification of mammographic parenchymal patterns. In *Medical Image Understanding and Analysis*, pages 145–148, 2001.

L Bocchi and J Nori. Shape analysis of microcalcifications using radon transform. *Medical Engineering & Physics*, 29(6):691–698, 2007.

L Bocchi, G Coppini, J Nori, and G Valli. Detection of single and clustered microcalcifications in mammograms using fractals models and neural networks. *Medical Engineering & Physics*, 26(4):303–312, 2004.

François O Bochud, Craig K Abbey, and Miguel P Eckstein. Statistical texture synthesis of mammographic images with clustered lumpy backgrounds. *Optics Express*, 4(1):33–43, 1999.

Carol A Bodian, Karl H Perzin, and Raffaele Lattes. Lobular neoplasia: long term risk of breast cancer and relation to other factors. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 78(5):1024–1034, 1996.

Anna Bosch, Xavier Munoz, Arnau Oliver, and Joan Marti. Modeling and classifying breast tissue density in mammograms. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, volume 2, pages 1552–1558. IEEE, 2006.

Alan C. Bovik, Marianna Clark, and Wilson S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (1):55–73, 1990.

Keir Bovis and Sameer Singh. Classification of mammographic breast density using a combined classifier paradigm. In *4th International Workshop on Digital Mammography*, pages 177–180, 2002.

NF Boyd, B O'sullivan, JE Campbell, E Fishell, I Simor, G Cooke, and T Germanson. Mammographic signs as risk factors for breast cancer. *British Journal of Cancer*, 45(2):185, 1982.

NF Boyd, JW Byng, RA Jong, EK Fishell, LE Little, AB Miller, GA Lockwood, DL Tritchler, and Martin J Yaffe. Quantitative classification of mammographic densities and breast cancer risk: results from the canadian national breast screening study. *JNCI: Journal of the National Cancer Institute*, 87(9):670–675, 1995.

Norman Boyd, Lisa Martin, Anoma Gunasekara, Olga Melnichouk, Gord Maudsley, Chris Peressotti, Martin Yaffe, and Salomon Minkin. Mammographic density and breast cancer risk: evaluation of a novel method of measuring breast tissue volumes. *Cancer Epidemiology and Prevention Biomarkers*, 18(6):1754–1762, 2009.

Norman F Boyd, Gina A Lockwood, Jeff W Byng, David L Tritchler, and Martin J Yaffe. Mammographic densities and breast cancer risk. *Cancer Epidemiology and Prevention Biomarkers*, 7(12):1133–1144, 1998.

Norman F Boyd, Helen Guo, Lisa J Martin, Limei Sun, Jennifer Stone, Eve Fishell, Roberta A Jong, Greg Hislop, Anna Chiarelli, Salomon Minkin, et al. Mammographic density and the risk and detection of breast cancer. *New England Journal of Medicine*, 356(3):227–236, 2007.

Roger Boyle, V Hlavac, and M Sonka. Image processing. *Analysis, and Machine Vision: Thomson Leaning*, 2008.

Andrew P Bradley. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145–1159, 1997.

Rachel F Brem, Janet Baum, Mary Lechner, Stuart Kaplan, Stuart Souders, L Gill Naul, and Jeff Hoffmeister. Improvement in sensitivity of screening mammography with computer-aided detection: a multiinstitutional trial. *American Journal of Roentgenology*, 181(3):687–693, 2003.

Rachel F Brem, Jeffrey W Hoffmeister, Jocelyn A Rapelyea, Gilat Zisman, Kevin Mohtashemi, Guarav Jindal, Martin P DiSimio, and Steven K Rogers. Impact of

breast density on computer-aided detection for breast cancer. *American Journal of Roentgenology*, 184(2):439–444, 2005a.

Rachel F Brem, Jeffrey W Hoffmeister, Gilat Zisman, Martin P DeSimio, and Steven K Rogers. A computer-aided detection system for the evaluation of breast cancer by mammographic appearance and lesion size. *American Journal of Roentgenology*, 184(3):893–896, 2005b.

Noel T Brewer, Talya Salz, and Sarah E Lillie. Systematic review: the long-term effects of false-positive mammograms. *Annals of Internal Medicine*, 146(7):502–510, 2007.

E Oran Brigham and CK Yuen. The fast fourier transform. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(2):146–146, 1978.

Jacques Brisson. Family history of breast cancer, mammographic features of breast tissue, and breast cancer risk. *Epidemiology*, pages 440–444, 1991.

Lawrence C Brody and Barbara Bowles Biesecker. Breast cancer susceptibility genes. brca1 and brca2. *Medicine*, 77(3):208–226, 1998.

Guanxiong Cai, Yanhui Guo, Yaqin Zhang, Genggeng Qin, Yuanpin Zhou, and Yao Lu. A fully automatic microcalcification detection approach based on deep convolution neural network. In *Medical Imaging 2018: Computer-Aided Diagnosis*, volume 10575, page 105752Q. International Society for Optics and Photonics, 2018a.

Sheng Cai, Jia-Neng Tang, Pei-Zhong Liu, Ya-Li Cai, Yan-Min Luo, Yong-Zhao Du, Yan Peng, and Ping Li. Breast density classification based on wavelet transform. *Journal of Medical Imaging and Health Informatics*, 8(6):1157–1163, 2018b.

Maria Julia Gregorio Calas, Bianca Gutfilen, and Wagner Coelho de Albuquerque Pereira. Cad and mammography: why use this tool? *Radiologia Brasileira*, 45(1): 46–52, 2012.

John Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6):679–698, 1986.

Patricia A Carney, Diana L Miglioretti, Bonnie C Yankaskas, Karla Kerlikowske, Robert Rosenberg, Carolyn M Rutter, Berta M Geller, Linn A Abraham, Steven H Taplin, Mark Dignan, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Annals of Internal Medicine*, 138(3):168–175, 2003.

Cyril Castella. Breast texture synthesis and estimation of the role of the anatomy and tumor shape in the radiological detection process. Technical report, EPFL, 2009.

Ronald A Castellino. Computer aided detection (cad): an overview. *Cancer Imaging*, 5(1):17, 2005.

Heang-Ping Chan, Kunio Doi, Carl J Vyborny, Robert A Schmidt, Charles E Metz, Kwok Leung Lam, Toshihiro Ogura, YZ Wu, and Heber MacMahon. Improvement in radiologists' detection of clustered microcalcifications on mammograms. the potential of computer-aided diagnosis. *Investigative Radiology*, 25(10):1102–1110, 1990.

Heang-Ping Chan, Berkman Sahiner, Nicholas Petrick, Mark A Helvie, Kwok Leung Lam, Dorit D Adler, and Mitchell M Goodsitt. Computerized classification of malignant and benign microcalcifications on mammograms: texture analysis using an artificial neural network. *Physics in Medicine & Biology*, 42(3):549, 1997.

Heang-Ping Chan, Berkman Sahiner, Kwok Leung Lam, Nicholas Petrick, Mark A Helvie, Mitchell M Goodsitt, and Dorit D Adler. Computerized analysis of mammographic microcalcifications in morphological and texture feature spaces. *Medical Physics*, 25(10):2007–2019, 1998.

Tianhorng Chang and C-CJ Kuo. Texture analysis and classification with tree-structured wavelet transform. *IEEE Transactions on Image Processing*, 2(4):429–441, 1993.

Sevastianos Chatzistergos, John Stoitsis, Angeliki Papaevangelou, Georgios Zografos, and Konstantina S Nikita. Parenchymal breast density estimation with the use of statistical characteristics and textons. In *Proceedings of the 10th IEEE International Conference on Information Technology and Applications in Biomedicine*, pages 1–4. IEEE, 2010.

Bidyut Baran Chaudhuri and Nirupam Sarkar. Texture segmentation using fractal dimension. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(1): 72–77, 1995.

Zhili Chen and Reyer Zwiggelaar. Segmentation of the breast region with pectoral muscle removal in mammograms. *Medical Image Understanding and Analysis (MIUA)*, pages 71–76, 2010.

Zhili Chen, Erika Denton, and Reyer Zwiggelaar. Local feature based mammographic tissue pattern modelling and breast density classification. In *Biomedical Engineering and Informatics (BMEI), 2011 4th International Conference on*, volume 1, pages 351–355. IEEE, 2011.

Zhili Chen, Arnau Oliver, Erika Denton, and Reyer Zwiggelaar. Automated mammographic risk classification based on breast density estimation. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 237–244. Springer, 2013a.

Zhili Chen, Liping Wang, Erika Denton, and Reyer Zwiggelaar. A multiscale blob representation of mammographic parenchymal patterns and mammographic risk assessment. In *International Conference on Computer Analysis of Images and Patterns*, pages 346–353. Springer, 2013b.

Zhili Chen, Harry Strange, Erika Denton, and Reyer Zwiggelaar. Analysis of mammographic microcalcification clusters using topological features. In *International Workshop on Digital Mammography*, pages 620–627. Springer, 2014.

Zhili Chen, Harry Strange, Arnau Oliver, Erika RE Denton, Caroline Boggis, and Reyer Zwiggelaar. Topological modeling and classification of mammographic microcalcification clusters. *IEEE Transactions on Biomedical Engineering*, 62(4): 1203–1214, 2015.

Heng-Da Cheng, Xiaopeng Cai, Xiaowei Chen, Liming Hu, and Xueling Lou. Computer-aided detection and classification of microcalcifications in mammograms: a survey. *Pattern Recognition*, 36(12):2967–2991, 2003.

Yiu-ming Cheung and Junping Deng. Ultra local binary pattern for image texture analysis. In *Security, Pattern Analysis, and Cybernetics (SPAC), 2014 International Conference on*, pages 290–293. IEEE, 2014.

Yateen Chitre, Atam P Dhawan, and Myron Moskowitz. Artificial neural network based classification of mammographic microcalcifications using image structure features. *International Journal of Pattern Recognition and Artificial Intelligence*, 7 (06):1377–1401, 1993.

S Ciatto, M Rosselli Del Turco, P Burke, C Visioli, E Paci, and M Zappa. Comparison of standard and double reading and computer-aided detection (cad) of interval cancers at prior negative screening mammograms: blind review. *British Journal of Cancer*, 89(9):1645, 2003.

Marcin Ciecholewski. Microcalcification segmentation from mammograms: A morphological approach. *Journal of Digital Imaging*, 30(2):172–184, 2017.

Elizabeth B Claus, Neil Risch, and W Douglas Thompson. Autosomal dominant inheritance of early-onset breast cancer. implications for risk prediction. *Cancer*, 73 (3):643–651, 1994.

David A Clausi. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing*, 28(1):45–62, 2002.

Fernand S. Cohen, Zhigang Fan, and Maqbool A Patel. Classification of rotated and scaled textured images using gaussian markov random field models. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (2):192–202, 1991.

Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (5): 603–619, 2002.

René Aloísio da Costa Vieira, Gabriele Biller, Gilberto Uemura, Carlos Alberto Ruiz, and Maria Paula Curado. Breast cancer screening in developing countries. *Clinics*, 72(4):244–253, 2017.

John S DaPonte and Porter Sherman. Classification of ultrasonic image texture by statistical discriminant analysis and neural networks. *Computerized Medical Imaging and Graphics*, 15(1):3–9, 1991.

DH Davies. Digital mammography—the comparative evaluation of film digitizers. *The British Journal of Radiology*, 66(790):930–933, 1993.

Rianne De Gelder, Eveline AM Heijnsdijk, Nicolien T Van Ravesteyn, Jacques Fracheboud, Gerrit Draisma, and Harry J De Koning. Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiologic Reviews*, 33(1):111–121, 2011.

BL De Stavola, IH Gravelle, DY Wang, DS Allen, RD Bulbrook, IS Fentiman, JL Hayward, and MC Chaudary. Relationship of mammographic parenchymal patterns with breast cancer risk factors and risk of breast cancer in a prospective study. *International Journal of Epidemiology*, 19(2):247–254, 1990.

Haluk Derin and William S Cole. Segmentation of textured images using gibbs random fields. *Computer Vision, Graphics, and Image Processing*, 35(1):72–98, 1986.

Stamatia Destounis, Sarah Hanson, Renee Morgan, Philip Murphy, Patricia Somerville, Posy Seifert, Valerie Andolina, Andrea Arieno, Melissa Skolny, and Wende Logan-Young. Computer-aided detection of breast carcinoma in standard mammographic projections with digital mammography. *International Journal of Computer Assisted Radiology and Surgery*, 4(4):331–336, 2009.

Stamatia V Destounis, Patricia DiNitto, Wende Logan-Young, Ermelinda Bonaccio, Margarita L Zuley, and Kathleen M Willison. Can computer-aided detection with double reading of screening mammograms help decrease the false-negative rate? initial experience. *Radiology*, 232(2):578–584, 2004.

Atam P Dhawan, Yateen Chitre, and Christine Kaiser-Bonasso. Analysis of mammographic microcalcifications using gray-level image structure features. *IEEE Transactions on Medical Imaging*, 15(3):246–259, 1996.

J Dheeba and S Tamil Selvi. Classification of malignant and benign microcalcification using svm classifier. In *Emerging Trends in Electrical and Computer Technology (ICETECT)*, pages 686–690. IEEE, 2011.

Tiago A Docusse, Jullyene R Furlani, Rodolfo P Romano, Rodrigo C Guido, Shi-Huang Chen, Norian Marranghello, and Aledir S Pereira. Microcalcification enhancement and classification on mammograms using the wavelet transform. In *Neural Networks, 2008. IJCNN 2008.(IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*, pages 3181–3186. IEEE, 2008.

Kunio Doi. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, 31(4-5): 198–211, 2007.

Alfonso Rojas Dominguez and Asoke K Nandi. Enhanced multi-level thresholding segmentation and rank based region selection for detection of masses in mammograms. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, volume 1, pages I–449. IEEE, 2007.

Carl J D'Orsi. *ACR BI-RADS atlas: breast imaging reporting and data system*. American College of Radiology, 2013.

Richard O Duda and Peter E Hart. Pattern classification and scene analysis. *A Wiley-Interscience Publication, New York: Wiley, 1973*, 1973.

Richard O Duda, Peter E Hart, and David G Stork. Pattern classification second edition john wiley & sons. *New York*, 58:16, 2001.

Richard O Duda, Peter E Hart, and David G Stork. *Pattern Classification*. John Wiley & Sons, 2012.

RPW Duin, J Kittler, M Hatef, and J Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239, 1998.

M D'Aloia, M Rizzi, and PD Bari. A second opinion system for microcalcification diagnosis. *World Appl. Sci. J*, 23(3):289–295, 2013.

CJ D'orsi, LW Bassett, S Feig, et al. Breast imaging reporting and data system (bi-rads). *Breast imaging atlas, 4th edn. American College of Radiology, Reston Google Scholar*, 1998.

Margaret M Eberl, Chester H Fox, Stephen B Edge, Cathleen A Carter, and Martin C Mahoney. Bi-rads classification for management of abnormal mammograms. *The Journal of the American Board of Family Medicine*, 19(2):161–164, 2006.

Suzanne A Eccles, Eric O Aboagye, Simak Ali, Annie S Anderson, Jo Armes, Fedor Berditchevski, Jeremy P Blaydes, Keith Brennan, Nicola J Brown, Helen E Bryant, et al. Critical research gaps and translational priorities for the successful prevention and treatment of breast cancer. *Breast Cancer Research*, 15(5):R92, 2013.

Alexei A Efros and William T Freeman. Image quilting for texture synthesis and transfer. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 341–346. ACM, 2001.

GW Eklund. The art of mammographic positioning. In *Radiological Diagnosis of Breast Diseases*, pages 75–88. Springer, 2000.

Ali El-Zaart. Expectation–maximization technique for fibro-glandular discs detection in mammography images. *Computers in Biology and Medicine*, 40(4):392–401, 2010.

Joann G Elmore, Carolyn K Wells, Carol H Lee, Debra H Howard, and Alvan R Feinstein. Variability in radiologists' interpretations of mammograms. *New England Journal of Medicine*, 331(22):1493–1499, 1994.

Joann G Elmore, Connie Y Nakano, Thomas D Koepsell, Laurel M Desnick, Carl J D'orsi, and David F Ransohoff. International variation in screening mammography interpretations in community-based programs. *Journal of the National Cancer Institute*, 95(18):1384–1393, 2003.

Joann G Elmore, Katrina Armstrong, Constance D Lehman, and Suzanne W Fletcher. Screening for breast cancer. *Jama*, 293(10):1245–1256, 2005.

Matthias Elter and Alexander Horsch. Cadx of mammographic masses and clustered microcalcifications: a review. *Medical Physics*, 36(6Part1):2052–2068, 2009.

Cynthia Y Enderwick and Ecangelia Micheli-Tzanakou. Classification of mammographic tissue using shape and texture features. In *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.'Magnificent Milestones and Emerging Opportunities in Medical Engineering'(Cat. No. 97CH36136)*, volume 2, pages 810–813. IEEE, 1997.

Kjersti Engan, Karl Skretting, Jostein Herredsvela, and Thor Ole Gulsrud. Frame texture classification method (ftcm) applied on mammograms for detection of abnormalities. *WASET International Journal of Signal Processing (IJSP)*, 4(2), 2007.

Rachida Es-salhi, Imane Daoudi, Saida Tallal, Hicham Medromi, et al. A survey on segmentation techniques of mammogram images. In *Advances in Ubiquitous Networking 2*, pages 545–556. Springer, 2017.

Nilofar Farzana and S Rathnakara. Detection of micro-calcifications in mammogram images using probabilistic neural network. *Elixir Comput. Sci. Eng*, 38:4404–4408, 2011.

Amir Fazlollahi, Fabrice Meriaudeau, Victor L Villemagne, Christopher C Rowe, Patricia M Desmond, Paul A Yates, Olivier Salvado, and Pierrick Bourgeat. Automatic detection of small spherical lesions using multiscale approach in 3d medical images. In *2013 IEEE International Conference on Image Processing*, pages 1158–1162. IEEE, 2013.

SA Feig, BM Galkin, and HD Muir. Evaluation of breast microcalcifications by means of optically magnified tissue specimen radiographs. In *Breast Cancer*, pages 111–123. Springer, 1987.

Stephen A Feig and R Edward Hendrick. Radiation risk from screening mammography of women aged 40-49 years. *JNCI Monographs*, 1997(22):119–124, 1997.

Stephen A Feig, Gary S Shaber, Gordon F Schwartz, Arthur Patchefsky, Herman I Libshitz, Jack Edeiken, Rudolph Nerlinger, Robert F Curley, and John D Wallace. Thermography, mammography, and clinical examination in breast cancer screening: review of 16,000 studies. *Radiology*, 122(1):123–127, 1977.

David P Feldman and James P Crutchfield. Structural information in two-dimensional patterns: Entropy convergence and excess entropy. *Physical Review E*, 67(5):051104, 2003.

Jaques Ferlay, Eva Steliarova-Foucher, Joannie Lortet-Tieulent, Sonia Rosso, Jan-Willem W Coebergh, Harry Comber, David Forman, and Freddie Bray. Cancer incidence and mortality patterns in europe: estimates for 40 countries in 2012. *European Journal of Cancer*, 49(6):1374–1403, 2013.

RJ Ferrari, RM Rangayyan, RA Borges, and AF Frere. Segmentation of the fibroglandular disc in mammogrms using gaussian mixture modelling. *Medical and Biological Engineering and Computing*, 42(3):378–387, 2004.

Pasquale Foggia, Carlo Sansone, Francesco Tortorella, and Mario Vento. Automatic classification of clustered microcalcifications by a multiple classifier system. In *International Workshop on Multiple Classifier Systems*, pages 208–217. Springer, 2001.

Erin EE Fowler, Celine M Vachon, Christopher G Scott, Thomas A Sellers, and John J Heine. Automated percentage of breast density measurements for full-field digital mammography applications. *Academic Radiology*, 21(8):958–970, 2014.

Timothy W Freer and Michael J Ulissey. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology*, 220(3):781–786, 2001.

Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *The annals of statistics*, 28(2):337–407, 2000.

Nir Friedman, Dan Geiger, and Moises Goldszmidt. Bayesian network classifiers. *Machine Learning*, 29(2):131–163, 1997.

JC Fu, SK Lee, STC Wong, JY Yeh, AH Wang, and HK Wu. Image segmentation feature selection and pattern classification for mammographic microcalcifications. *Computerized Medical Imaging and Graphics*, 29(6):419–429, 2005.

Mitchell H Gail, Louise A Brinton, David P Byar, Donald K Corle, Sylvan B Green, Catherine Schairer, and John J Mulvihill. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *JNCI: Journal of the National Cancer Institute*, 81(24):1879–1886, 1989.

Csaba Gajdos, Paul Ian Tartter, Ira J Bleiweiss, George Hermann, John De Csepel, Alison Estabrook, and Alfred W Rademaker. Mammographic appearance of nonpalpable breast cancer reflects pathologic characteristics. *Annals of Surgery*, 235(2): 246, 2002.

Mary M Galloway. Texture analysis using grey level run lengths. *NASA STI/Recon Technical Report N*, 75, 1974.

Aimilia Gastounioti, Emily F Conant, and Despina Kontos. Beyond breast density: a review on the advancing role of parenchymal texture analysis in breast cancer risk assessment. *Breast Cancer Research*, 18(1):91, 2016.

Shantanu Gaur, Vandana Dialani, Priscilla J Slanetz, and Ronald L Eisenberg. Architectural distortion of the breast. *American Journal of Roentgenology*, 201(5): W662–W670, 2013.

K Geetha, K Thanushkodi, et al. New particle swarm optimization for feature selection and classification of microcalcifications in mammograms. In *2008 International Conference on Signal Processing, Communications and Networking*, pages 458–463. IEEE, 2008.

Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6):721–741, 1984.

M George, ER Denton, and R Zwiggelaar. Topological connected chain modelling for classification of mammographic microcalcification. In *Proceedings of the Conference on Computer Graphics & Visual Computing*, pages 1–5. Eurographics Association, 2018a.

Minu George, Erika Denton, and Reyer Zwiggelaar. Mammogram breast density classification using mean-elliptical local binary patterns. In *14th International Workshop on Breast Imaging (IWBI 2018)*, volume 10718, page 107180B. International Society for Optics and Photonics, 2018b.

Minu George, Erika R. E. Denton, and Reyer Zwiggelaar. Topological Connected Chain Modelling for Classification of Mammographic Microcalcification. In *Computer Graphics and Visual Computing (CGVC)*. The Eurographics Association, 2018c. ISBN 978-3-03868-071-0.

Jeffrey S Geronimo, Douglas P Hardin, and Peter R Massopust. Fractal functions and wavelet expansions based on several scaling functions. *Journal of Approximation Theory*, 78(3):373–401, 1994.

David J Getty, Ronald M Pickett, Carl J D'orsi, and John A Swets. 18 enhanced interpretation of diagnostic images. *Judgment and Decision Making: An Interdisciplinary Reader*, page 281, 2000.

James J Gibson. The perception of the visual world. 1950.

Fiona J Gilbert, Susan M Astley, Magnus A McGee, Maureen GC Gillan, Caroline RM Boggis, Pamela M Griffiths, and Stephen W Duffy. Single reading with computer-aided detection and double reading of screening mammograms in the united kingdom national breast screening program. *Radiology*, 241(1):47–53, 2006.

Peter C Gøtzsche, Margrethe Nielsen, et al. Screening for breast cancer with mammography. *Cochrane Database Syst Rev*, 4(1), 2009.

Jon M Greif. Mammographic screening for breast cancer: An invited review of the benefits and costs. *The Breast*, 19(4):268–272, 2010.

Matthew Gromet. Comparison of computer-aided detection to double reading of screening mammograms: review of 231,221 mammograms. *American Journal of Roentgenology*, 190(4):854–859, 2008.

Zhenhua Guo, Lei Zhang, and David Zhang. Rotation invariant texture classification using lbp variance (lbpv) with global matching. *Pattern Recognition*, 43(3):706–719, 2010a.

Zhenhua Guo, Lei Zhang, David Zhang, and Su Zhang. Rotation invariant texture classification using adaptive lbp with directional statistical features. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 285–288. IEEE, 2010b.

Lubomir Hadjiiski, Berkman Sahiner, and Heang-Ping Chan. Advances in cad for diagnosis of breast cancer. *Current Opinion in Obstetrics & Gynecology*, 18(1):64, 2006.

Edward M Hadley, Erika RE Denton, Josep Pont, Elsa Pérez, and Reyer Zwiggelaar. Analysis of anatomical linear structure information in mammographic risk assessment. In *International Workshop on Digital Mammography*, pages 483–490. Springer, 2008.

George M Haley and BS Manjunath. Rotation-invariant texture classification using a complete space-frequency model. *IEEE Transactions on Image Processing*, 8(2): 255–269, 1999.

Mark Andrew Hall. Correlation-based feature selection for machine learning. 1999.

Mark D Halling-Brown, PT Looney, MN Patel, LM Warren, A Mackenzie, and KC Young. The oncology medical image database (omi-db). In *Medical Imaging 2014: PACS and Imaging Informatics: Next Generation and Innovations*, volume 9039, page 903906. International Society for Optics and Photonics, 2014.

Nezha Hamdi, Khalid Auhmani, and Moha M'rabet Hassani. Design of a high-accuracy classifier based on fisher discriminant analysis: Application to computer-aided diagnosis of microcalcifications. In *2008 International Conference on Computational Sciences and Its Applications*, pages 267–273. IEEE, 2008.

James A Hanley, Ailish Hannigan, and Katie M O'Brien. Mortality reductions due to mammography screening: Contemporary population-based data. *PloS one*, 12(12): e0188947, 2017.

Robert M Haralick. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786–804, 1979.

Robert M Haralick, K Shanmugam, Its'Hak Dinstein, et al. Textural features for image classification. *IEEE Transactions on systems, Man, and Cybernetics*, 3(6):610–621, 1973.

Chen Chi Hau. *Handbook of pattern recognition and computer vision*. World Scientific, 2015.

Chu He, Timo Ahonen, and Matti Pietikainen. A bayesian local binary pattern texture descriptor. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2008a.

Wenda He and Reyer Zwiggelaar. Breast parenchymal pattern analysis in digital mammography: associations between tabár and birads tissue compositions. In *International Conference on Computer Analysis of Images and Patterns*, pages 386–393. Springer, 2013.

Wenda He, Izzati Muhimmah, Erika RE Denton, and Reyer Zwiggelaar. Mammographic segmentation based on texture modelling of tabár mammographic building blocks. In *International Workshop on Digital Mammography*, pages 17–24. Springer, 2008b.

Wenda He, Erika RE Denton, and Reyer Zwiggelaar. Mammographic segmentation based on mammographic parenchymal patterns and spatial moments. In *2009 9th International Conference on Information Technology and Applications in Biomedicine*, pages 1–4. IEEE, 2009.

Wenda He, Erika RE Denton, and Reyer Zwiggelaar. Mammographic image segmentation and risk classification using a novel texture signature based methodology. In *International Workshop on Digital Mammography*, pages 526–533. Springer, 2010a.

Wenda He, Erika RE Denton, and Reyer Zwiggelaar. Mammographic segmentation and risk classification using a novel binary model based bayes classifier. In *International Workshop on Digital Mammography*, pages 40–47. Springer, 2012.

Wenda He, Arne Juette, Erika RE Denton, Arnau Oliver, Robert Martí, and Reyer Zwiggelaar. A review on automatic mammographic density and parenchymal segmentation. *International Journal of Breast Cancer*, 2015, 2015.

Yonggang He, Nong Sang, and Changxin Gao. Pyramid-based multi-structure local binary pattern for texture classification. In *Asian Conference on Computer Vision*, pages 133–144. Springer, 2010b.

M Heath, K Bowyer, D Kopans, R Moore, and P Kegelmeyer. The digital database for screening mammography. *Digital Mammography*, pages 431–434, 2000.

John J Heine, Stanley R Deans, Robert P Velthuizen, and Laurence P Clarke. On the statistical nature of mammograms. *Medical Physics*, 26(11):2254–2265, 1999.

P Henrot, A Leroux, C Barlier, and P Génin. Breast microcalcifications: the lesions in anatomical pathology. *Diagnostic and Interventional Imaging*, 95(2):141–152, 2014.

Patricia Lorena Arancibia Hernández, Teresa Taub Estrada, Alejandra López Pizarro, María Lorena Díaz Cisternas, and C Sáez Tapia. Breast calcifications: description and classification according to bi-rads 5th edition. *Rev. Chil. Radiol.*, 22:80–91, 2016.

Christian Herrmann, Penelope Vounatsou, Beat Thürlimann, Nicole Probst-Hensch, Christian Rothermundt, and Silvia Ess. Impact of mammography screening programmes on breast cancer mortality in switzerland, a country with different regional screening policies. *BMJ open*, 8(3):e017806, 2018.

K Nyem Nyem Hlaing. First order statistics and glcm based feature extraction for recognition of myanmar paper currency. In *Proceedings of the 31st IIER International Conference, Bangkok, Thailand*, pages 1–6, 2015.

WT Ho and PWT Lam. Clinical performance of computer-assisted detection (cad) system in detecting carcinoma in breasts of different densities. *Clinical Radiology*, 58(2):133–136, 2003.

Roland Holland and JH Hendriks. Microcalcifications associated with ductal carcinoma in situ: mammographic-pathologic correlation. In *Seminars in Diagnostic Pathology*, volume 11, pages 181–192, 1994.

Karla Horsch, Maryellen L Giger, Carl J Vyborny, Li Lan, Ellen B Mendelson, and R Edward Hendrick. Classification of breast lesions with multimodality computer-aided diagnosis: observer study results on an independent clinical data set. *Radiology*, 240(2):357–368, 2006.

Jianping Hua, Zixiang Xiong, James Lowey, Edward Suh, and Edward R Dougherty. Optimal number of features as a function of sample size for various classification rules. *Bioinformatics*, 21(8):1509–1515, 2005.

Gordon Hughes. On the mean accuracy of statistical pattern recognizers. *IEEE transactions on information theory*, 14(1):55–63, 1968.

Anne Humeau-Heurtier. Texture feature extraction methods: A survey. *IEEE Access*, 7:8975–9000, 2019.

Zhimin Huo, Maryellen L Giger, Carl J Vyborny, and Charles E Metz. Breast cancer: effectiveness of computer-aided diagnosis—observer study with independent database of mammograms. *Radiology*, 224(2):560–568, 2002.

Sibt Ul Hussain, Thibault Napoléon, and Fréderic Jurie. Face recognition using local quantized patterns. In *British Machine Vision Conference*, pages 11–pages, 2012.

Dimitris K Iakovidis, Eystratios G Keramidas, and Dimitris Maroulis. Fuzzy local binary patterns for ultrasound texture characterization. In *International conference image analysis and recognition*, pages 750–759. Springer, 2008.

Mohammad Ibrahim, Md Iftekharul Alam Efat, Humayun Kayesh Shamol, Shah Mostafa Khaled, Mohammad Shoyaib, and M Abdullah-Al-Wadud. Dynamic local ternary pattern for face recognition and verification. In *Proceedings of the International Conference on Computer Engineering and Applications, Tenerife, Spain*, volume 1012, 2014.

Taskeed Jabid, Md Hasanul Kabir, and Oksam Chae. Local directional pattern (ldp) for face recognition. In *Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on*, pages 329–330. IEEE, 2010.

Kourosh Jafari-Khouzani and Hamid Soltanian-Zadeh. Rotation-invariant multiresolution texture analysis using radon and wavelet transforms. *IEEE Transactions on Image Processing*, 14(6):783–795, 2005.

Anil K Jain, M Narasimha Murty, and Patrick J Flynn. Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3):264–323, 1999.

Ahmedin Jemal, Rebecca Siegel, Elizabeth Ward, Yongping Hao, Jiaquan Xu, and Michael J Thun. Cancer statistics, 2009. *CA: A Cancer Journal for Clinicians*, 59 (4):225–249, 2009.

Richard Jensen and Qiang Shen. Are more features better? a response to attributes reduction using fuzzy rough sets. *IEEE Transactions on Fuzzy Systems*, 17(6): 1456–1458, 2009.

Pavel Jetensky. Human hand image analysis extracting finger coordinates using circular scanning. In *2015 25th International Conference Radioelektronika (RADIOELEKTRONIKA)*, pages 427–430. IEEE, 2015.

Tao Jia, Nong-liang Sun, and Mao-yong Cao. Moving object detection based on blob analysis. In *2008 IEEE International Conference on Automation and Logistics*, pages 322–325. IEEE, 2008.

Yulei Jiang, Robert M Nishikawa, Robert A Schmidt, Alicia Y Toledano, and Kunio Doi. Potential of computer-aided diagnosis to reduce variability in radiologists' interpretations of mammograms depicting microcalcifications. *Radiology*, 220(3): 787–794, 2001.

Hongliang Jin, Qingshan Liu, Hanqing Lu, and Xiaofeng Tong. Face detection using improved lbp under bayesian framework. In *Image and Graphics (ICIG'04), Third International Conference on*, pages 306–309. IEEE, 2004.

Nabila Kadaoui, Maryse Guay, Geneviève Baron, José St-Cerny, and Jacques Lemaire. Breast cancer screening practices for women aged 35 to 49 and 70 and older. *Canadian Family Physician*, 58(1):e47–e53, 2012.

Michiel Kallenberg, Kersten Petersen, Mads Nielsen, Andrew Y Ng, Pengfei Diao, Christian Igel, Celine M Vachon, Katharina Holland, Rikke Rass Winkel, Nico Karssemeijer, et al. Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring. *IEEE Transactions on Medical Imaging*, 35 (5):1322–1331, 2016.

Maria Kallergi. Computer-aided diagnosis of mammographic microcalcification clusters. *Medical Physics*, 31(2):314–326, 2004.

Farin Kamangar, Graça M Dores, and William F Anderson. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *Journal of Clinical Oncology*, 24(14):2137–2150, 2006.

Marzena Kamińska, Tomasz Ciszewski, Karolina Łopacka-Szatan, Paweł Miotła, and Elżbieta Starosławska. Breast cancer risk factors. *Przeglad menopauzalny= Menopause review*, 14(3):196, 2015.

Lance M Kaplan. Extended fractal analysis for texture classification and segmentation. *IEEE Transactions on Image Processing*, 8(11):1572–1585, 1999.

Murat Karabatak. A new classifier for breast cancer detection based on naïve bayesian. *Measurement*, 72:32–36, 2015.

A Karahaliou, S Skiadopoulos, I Boniatis, P Sakellaropoulos, E Likaki, G Panayiotakis, and L Costaridou. Texture analysis of tissue surrounding microcalcifications on mammograms for breast cancer diagnosis. *The British Journal of Radiology*, 80 (956):648–656, 2007.

Anna N Karahaliou, Ioannis S Boniatis, Spyros G Skiadopoulos, Filippos N Sakellaropoulos, Nikolaos S Arikidis, Eleni A Likaki, George S Panayiotakis, and Lena I Costaridou. Breast cancer diagnosis: analyzing texture of tissue surrounding microcalcifications. *IEEE Transactions on Information Technology in Biomedicine*, 12(6): 731–738, 2008.

Nico Karssemeijer. A stochastic model for automated detection of calcifications in digital mammograms. In *Biennial International Conference on Information Processing in Medical Imaging*, pages 227–238. Springer, 1991.

Nico Karssemeijer. Automated classification of parenchymal patterns in mammograms. *Physics in Medicine & Biology*, 43(2):365, 1998.

Anne Kaspers. Blob detection. Master's thesis, 2011.

Tomi Kauppi, Joni-Kristian Kamarainen, Lasse Lensu, Valentina Kalesnykiene, Iiris Sorri, Heikki Kälviäinen, Hannu Uusitalo, and Juhani Pietilä. Fusion of multiple expert annotations and overall score selection for medical image diagnosis. In *Scandinavian Conference on Image Analysis*, pages 760–769. Springer, 2009.

K Kerlikowske, D Grady, and J Barclay. Effect of age, breast density and family history on the sensitivity of first screening. *JAMA*, 276:33–8, 1966.

Baljit Singh Khehra and Amar Partap Singh Pharwaha. Classification of clustered microcalcifications using mlffbp-ann and svm. *Egyptian Informatics Journal*, 17(1): 11–20, 2016.

Jong Kook Kim and Hyun Wook Park. Statistical textural features for detection of microcalcifications in digitized mammograms. *IEEE Transactions on Medical Imaging*, 18(3):231–238, 1999.

Youngwoo Kim and Jong Hyo Kim. Reliability of breast density estimation in follow-up mammograms: repeatability and reproducibility of a fully automated areal percent density method. In *International Workshop on Digital Mammography*, pages 304–311. Springer, 2014.

Youngwoo Kim, Changwon Kim, and Jong-Hyo Kim. Automated estimation of breast density on mammogram using combined information of histogram statistics and boundary gradients. In *Medical Imaging 2010: Computer-Aided Diagnosis*, volume 7624, page 76242F. International Society for Optics and Photonics, 2010.

Divya Kiran, Abdul Imran Rasheed, and Hariharan Ramasangu. Fpga implementation of blob detection algorithm for object detection in visual navigation. In *2013 International conference on Circuits, Controls and Communications (CCUBE)*, pages 1–5. IEEE, 2013.

Byoung Chul Ko, Seong Hoon Kim, and Jae-Yeal Nam. X-ray image classification using random forests with local wavelet-based cs-local binary patterns. *Journal of Digital Imaging*, 24(6):1141–1151, 2011.

Catherine M Kocur, Steven K Rogers, Lemuel R Myers, Thomas Burns, Matthew Kabrisky, Jeffrey W Hoffmeister, KW Bauer, and Jean M Steppe. Using neural networks to select wavelet features for breast cancer diagnosis. *IEEE Engineering in Medicine and Biology Magazine*, 15(3):95–102, 1996.

Jan J Koenderink. The structure of images. *Biological Cybernetics*, 50(5):363–370, 1984.

Thomas M Kolb, Jacob Lichy, and Jeffrey H Newhouse. Comparison of the performance of screening mammography, physical examination, and breast us and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology*, 225(1):165–175, 2002.

Hui Kong, Hatice Cinar Akakin, and Sanjay E Sarma. A generalized laplacian of gaussian filter for blob detection and its applications. *IEEE Transactions on Cybernetics*, 43(6):1719–1733, 2013a.

Hui Kong, Sanjay E Sarma, and Feng Tang. Generalizing laplacian of gaussian filters for vanishing-point detection. *IEEE Transactions on Intelligent Transportation Systems*, 14(1):408–418, 2013b.

D Kopans 3rd. Breast imaging 3rd ed, 2007.

Dani Kramer and Farzin Aghdasi. Classification of microcalcifications in digitised mammograms using multiscale statistical texture analysis. In *Proceedings of the 1998 South African Symposium on Communications and Signal Processing-COMSIG'98 (Cat. No. 98EX214)*, pages 121–126. IEEE, 1998.

Dani Kramer and Farzin Aghdasi. Texture analysis techniques for the classification of microcalcifications in digitised mammograms. In *1999 IEEE Africon. 5th Africon Conference in Africa (Cat. No. 99CH36342)*, volume 1, pages 395–400. IEEE, 1999.

S Krishnaveni, R Bhanumathi, and T Pugazharasan. Study of mammogram microcalcification to aid tumour detection using naive bayes classifier. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 3(3), 2014.

Indrajeet Kumar, Jitendra Virmani, and HS Bhadauria. A review of breast density classification methods. In *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 1960–1967. IEEE, 2015.

Vijay Kumar and Priyanka Gupta. Importance of statistical measures in digital image processing. *International Journal of Emerging Technology and Advanced Engineering*, 2(8):56–62, 2012.

SW-C Lam. Texture feature extraction using gray level gradient based co-occurence matrices. In *Systems, Man, and Cybernetics, 1996., IEEE International Conference on*, volume 1, pages 267–271. IEEE, 1996.

Vincenzo Lattanzio and Angela Maria Guerrieri. Mammographic breast density and its effects on imaging. In *Contrast-Enhanced Digital Mammography (CEDM)*, pages 9–16. Springer, 2018.

Kenneth I Laws. Texture energy measures. In *Proc. Image understanding Workshop*, pages 47–51, 1979.

Kenneth I Laws. Textured image segmentation. Technical report, University of Southern California Los Angeles Image Processing INST, 1980a.

KI Laws. Rapid texture identification. 24th annual technical symposium: International society for optics and photonics. 1980b.

Carol H Lee, D David Dershaw, Daniel Kopans, Phil Evans, Barbara Monsees, Debra Monticciolo, R James Brenner, Lawrence Bassett, Wendie Berg, Stephen Feig, et al. Breast cancer screening with imaging: recommendations from the society of breast imaging and the acr on the use of mammography, breast mri, breast ultrasound, and other technologies for the detection of clinically occult breast cancer. *Journal of the American college of Radiology*, 7(1):18–27, 2010.

Constance D Lehman, Robert D Wellman, Diana SM Buist, Karla Kerlikowske, Anna NA Tosteson, and Diana L Miglioretti. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Internal Medicine*, 175(11):1828–1837, 2015.

I Leichter, R Lederman, SS Buchbinder, P Bamberger, B Novak, and S Fields. Computerized evaluation of mammographic lesions: What diagnostic role does the shape of the individual microcalcifications play compared with the geometry of the cluster? *American Journal of Roentgenology*, 182(3):705–712, 2004.

Subhash Lele and Joan T Richtsmeier. Euclidean distance matrix analysis: A coordinate-free approach for comparing biological shapes using landmark data. *American Journal of Physical Anthropology*, 86(3):415–427, 1991.

RA Lerski, K Straughan, LR Schad, D Boyce, S Blüml, and I Zuna. Viii. mr image texture analysis—an approach to tissue characterization. *Magnetic Resonance Imaging*, 11(6):873–887, 1993.

Thomas Leung and Jitendra Malik. Representing and recognizing the visual appearance of materials using three-dimensional textons. *International Journal of Computer Vision*, 43(1):29–44, 2001.

Huai Li, KJ Ray Liu, and S-CB Lo. Fractal modeling and segmentation for the enhancement of microcalcifications in digital mammograms. *IEEE Transactions on Medical Imaging*, 16(6):785–798, 1997.

Hui Li, Maryellen L Giger, Zhimin Huo, Olufunmilayo I Olopade, Li Lan, Barbara L Weber, and Ioana Bonta. Computerized analysis of mammographic parenchymal patterns for assessing breast cancer risk: effect of roi size and location. *Medical Physics*, 31(3):549–555, 2004.

Jian Li, Qian Du, and Caixin Sun. An improved box-counting method for image fractal dimension estimation. *Pattern Recognition*, 42(11):2460–2469, 2009.

Shu Liao and Albert CS Chung. Face recognition by using elongated local binary patterns with average maximum distance gradient magnitude. In *Asian Conference on Computer Vision*, pages 672–679. Springer, 2007.

G Linda and C George Shapiro. Stockman, computer vision, 2001.

Tobias Lindahl. Study of local binary patterns, 2007.

Tony Lindeberg. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention. *International Journal of Computer Vision*, 11(3):283–318, 1993.

Tony Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):79–116, 1998.

Tony Lindeberg. Scale-space. *Wiley Encyclopedia of Computer Science and Engineering*, pages 2495–2504, 2007.

Tony Lindeberg. *Scale-space theory in computer vision*, volume 256. Springer Science & Business Media, 2013.

Li Liu, Jian Wang, and Kai He. Breast density classification using histogram moments of multiple resolution mammograms. In *2010 3rd International Conference on Biomedical Engineering and Informatics*, volume 1, pages 146–149. IEEE, 2010.

Li Liu, Paul Fieguth, Yulan Guo, Xiaogang Wang, and Matti Pietikäinen. Local binary features for texture classification: taxonomy and experimental study. *Pattern Recognition*, 62:135–160, 2017.

Qingqing Liu, Li Liu, Yanli Tan, Jian Wang, Xueyun Ma, and Hairi Ni. Mammogram density estimation using sub-region classification. In *Biomedical Engineering and Informatics (BMEI), 2011 4th International Conference on*, volume 1, pages 356–359. IEEE, 2011.

Joseph Y Lo, Marios A Gavrielides, Mia K Markey, and Jonathan L Jesneck. Computer-aided classification of breast microcalcification clusters: Merging of features from image processing and radiologists. In *Medical Imaging 2003: Image Processing*, volume 5032, pages 882–890. International Society for Optics and Photonics, 2003.

Jiwen Lu, Venice Erin Liong, and Jie Zhou. Simultaneous local binary feature learning and encoding for homogeneous and heterogeneous face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(8):1979–1993, 2018.

Shin-Yee Lu and King Sun Fu. A syntactic approach to texture analysis. *Computer Graphics and Image Processing*, 7(3):303–330, 1978.

Caro Lucas and Gelareh Mortezaie. Analysis of mammographic microcalcifications using gray-level images and neural networks. 2002.

Yiming Ma, Peter C Tay, Robert D Adams, and James Z Zhang. A novel shape feature to classify microcalcifications. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 2265–2268. IEEE, 2010.

E Malar, A Kandaswamy, D Chakravarthy, and A Giri Dharan. A novel approach for detection and classification of mammographic microcalcifications using wavelet analysis and extreme learning machine. *Computers in Biology and Medicine*, 42(9): 898–905, 2012.

Stephane G Mallat. A theory for multiresolution signal decomposition: the wavelet representation. 1987.

Margaret T Mandelson, Nina Oestreicher, Peggy L Porter, Donna White, Charles A Finder, Stephen H Taplin, and Emily White. Breast density as a predictor of mammographic detection: comparison of interval-and screen-detected cancers. *Journal of the National Cancer Institute*, 92(13):1081–1087, 2000.

Katherine E Martin, Mark A Helvie, Chuan Zhou, Marilyn A Roubidoux, Janet E Bailey, Chintana Paramagul, Caroline E Blane, Katherine A Klein, Seema S Sonnad, and Heang-Ping Chan. Mammographic density measured with quantitative computer-aided method: comparison with radiologists' estimates and bi-rads categories. *Radiology*, 240(3):656–665, 2006.

Christiane Marx, Ansgar Malich, Mirjam Facius, Uta Grebenstein, Dieter Sauner, Stefan OR Pfleiderer, and Werner A Kaiser. Are unnecessary follow-up procedures induced by computer-aided diagnosis (cad) in mammography? comparison of mammographic diagnosis with and without use of cad. *European Journal of Radiology*, 51(1):66–72, 2004.

M Masek, SM Kwok, CJS deSilva, and Y Attikiouzel. Classification of mammographic density using histogram distance measures. In *Proceedings of the World Congress on Medical Physics and Biomedical Engineering, August 2003, Sydney, Australia*, volume 1, 2003.

Gertraud Maskarinec, Ian Pagano, Galina Lurie, Lynne R Wilkens, and Laurence N Kolonel. Mammographic density and breast cancer risk: the multiethnic cohort study. *American Journal of Epidemiology*, 162(8):743–752, 2005.

Andrzej Materka, Michal Strzelecki, et al. Texture analysis methods–a review. *Technical university of lodz, institute of electronics, COST B11 report, Brussels*, pages 9–11, 1998.

Tomoko Matsubara, D Yamazaki, Masaya Kato, Takeshi Hara, Hiroshi Fujita, Takuji Iwase, and Tokiko Endo. An automated classification scheme for mammograms based on amount and distribution of fibroglandular breast tissue density. In *International Congress Series*, volume 1230, pages 545–552. Elsevier, 2001.

Susan Mayor. Mammography screening reduces breast cancer deaths in ireland, study finds. *BMJ: British Medical Journal (Online)*, 359, 2017.

Valerie A McCormack and Isabel dos Santos Silva. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiology and Prevention Biomarkers*, 15(6):1159–1169, 2006.

Peter Miller and Sue Astley. Classification of breast tissue by texture analysis. *Image and Vision Computing*, 10(5):277–282, 1992.

Aly A Mohamed, Wendie A Berg, Hong Peng, Yahong Luo, Rachel C Jankowitz, and Shandong Wu. A deep learning method for classifying mammographic breast density categories. *Medical Physics*, 45(1):314–321, 2018.

Abubacker Kaja Mohideen and Kuttiannan Thangavel. Region-based contrast enhancement of digital mammograms using an improved watershed segmentation. *International Journal of Image and Graphics*, 13(01):1350007, 2013.

Lia Morra, Daniela Sacchetto, Manuela Durando, Silvano Agliozzo, Luca Alessandro Carbonaro, Silvia Delsanto, Barbara Pesce, Diego Persano, Giovanna Mariscotti, Vincenzo Marra, et al. Breast cancer: computer-aided detection with digital breast tomosynthesis. *Radiology*, 277(1):56–63, 2015.

William Mark Morrow, Raman Bhalachandra Paranjape, Rangaraj M Rangayyan, and JE Leo Desautels. Region-based contrast enhancement of mammograms. *IEEE Transactions on Medical Imaging*, 11(3):392–406, 1992.

Marilyn J Morton, Dana H Whaley, Kathleen R Brandt, and Kimberly K Amrami. Screening mammograms: interpretation with computer-aided detection—prospective evaluation. *Radiology*, 239(2):375–383, 2006.

Rafayah Mousa, Qutaishat Munib, and Abdallah Moussa. Breast cancer diagnosis system based on wavelet analysis and fuzzy-neural. *Expert systems with Applications*, 28(4):713–723, 2005.

I Muhimmah, W He, ERE Denton, and R Zwiggelaar. Segmentation based on textons and mammographic building blocks. In *Medical Image Understanding and Analysis*, volume 2007, pages 228–232. Citeseer, 2007.

Izzati Muhimmah and Reyer Zwiggelaar. Mammographic density classification using multiresolution histogram information. In *Proceedings of the International Special Topic Conference on Information Technology in Biomedicine, ITAB*, pages 26–28, 2006.

Izzati Muhimmah, Arnau Oliver, Erika RE Denton, Josep Pont, Elsa Pérez, and Reyer Zwiggelaar. Comparison between wolfe, boyd, bi-rads and tabár based mammographic risk assessment. In *International Workshop on Digital Mammography*, pages 407–415. Springer, 2006.

Ramakrishnan Mukundan. Image features based on characteristic curves and local binary patterns for automated her2 scoring. *Journal of Imaging*, 4(2):35, 2018.

Mario Mustra, Mislav Grgic, and Kresimir Delac. Enhancement of microcalcifications in digital mammograms. In *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 248–251. IEEE, 2012.

Mario Muštra, Mislav Grgić, and Krešimir Delač. Breast density classification using multiple feature selection. *Automatika*, 53(4):362–372, 2012.

M Muttarak, P Kongmebhol, and N Sukhamwang. Breast calcifications: which are malignant. *Singapore Medical Journal*, 50(9):907–914, 2009.

Abdullah-Al Nahid and Yinan Kong. Involvement of machine learning for breast cancer image classification: A survey. *Computational and Mathematical Methods in Medicine*, 2017, 2017.

Yojana V Nalawade. Evaluation of breast calcifications. *Indian Journal of Radiology & Imaging*, 19(4), 2009.

Loris Nanni, Alessandra Lumini, and Sheryl Brahnam. Local binary patterns variants as texture descriptors for medical image analysis. *Artificial Intelligence in Medicine*, 49(2):117–125, 2010.

Steven A Narod, Javaid Iqbal, and Anthony B Miller. Why have breast cancer mortality rates declined? *Journal of Cancer Policy*, 5:8–17, 2015.

JT Neyhart, M Kirlakovsky, LM Coleman, R Polikar, M Tseng, and SA Mandayam. Automated segmentation and quantitative characterization of radiodense tissue in digitized mammograms. In *AIP Conference Proceedings*, volume 615, pages 1866–1873. AIP, 2002.

Huu-Tuan Nguyen and Alice Caplier. Elliptical local binary patterns for face recognition. In *Asian Conference on Computer Vision*, pages 85–96. Springer, 2012.

Robert M Nishikawa. Current status and future directions of computer-aided diagnosis in mammography. *Computerized Medical Imaging and Graphics*, 31(4-5):224–235, 2007.

Silvia Obenauer, Christian Sohns, Carola Werner, and Eckhardt Grabbe. Impact of breast density on computer-aided detection in full-field digital mammography. *Journal of Digital Imaging*, 19(3):258, 2006.

T Ojala and M Pietikäinen. Texture classification. machine vision and media processing unit, university of oulu, finland, 2004.

Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51–59, 1996.

Timo Ojala, Matti Pietikäinen, and Topi Mäenpää. Gray scale and rotation invariant texture classification with local binary patterns. In *European Conference on Computer Vision*, pages 404–420. Springer, 2000.

Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

Arnau Oliver, Jordi Freixenet, Anna Bosch, David Raba, and Reyer Zwiggelaar. Automatic classification of breast tissue. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 431–438. Springer, 2005a.

Arnau Oliver, Jordi Freixenet, and Reyer Zwiggelaar. Automatic classification of breast density. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, volume 2, pages II–1258. IEEE, 2005b.

Arnau Oliver, Jordi Freixenet, Robert Martí, and Reyer Zwiggelaar. A comparison of breast tissue classification techniques. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 872–879. Springer, 2006.

Arnau Oliver, Xavier Lladó, R Marti, Jordi Freixenet, and Reyer Zwiggelaar. Classifying mammograms using texture information. In *Medical Image Understanding and Analysis*, volume 223. Citeseer, 2007.

Arnau Oliver, Jordi Freixenet, Robert Marti, Josep Pont, Elsa Pérez, Erika RE Denton, and Reyer Zwiggelaar. A novel breast tissue density classification methodology. *IEEE Transactions on Information Technology in Biomedicine*, 12(1):55–65, 2008.

Arnau Oliver, Xavier Lladó, Elsa Pérez, Josep Pont, Erika RE Denton, Jordi Freixenet, and Joan Martí. A statistical approach for breast density segmentation. *Journal of Digital Imaging*, 23(5):527–537, 2010.

Arnau Oliver, Albert Torrent, Xavier Lladó, Meritxell Tortajada, Lidia Tortajada, Melcior Sentís, Jordi Freixenet, and Reyer Zwiggelaar. Automatic microcalcification and cluster detection for digital and digitised mammograms. *Knowledge-Based Systems*, 28:68–75, 2012.

Arnau Oliver, i Malagelada, Meritxell Tortajada Giménez, Xavier Lladó Bardera, Jordi Freixenet i Bosch, Sergi Ganau, Lídia Tortajada, Mariona Vilagran, Melcior Sentís, and Robert Martí Marly. Breast density analysis using an automatic density segmentation algorithm. © *Journal of Digital Imaging, 2015, vol. 28, núm. 5, p. 604-612*, 2015.

Ketil Oppedal, Trygve Eftestøl, Kjersti Engan, Mona K Beyer, and Dag Aarsland. Classifying dementia using local binary patterns from different regions in magnetic resonance images. *Journal of Biomedical Imaging*, 2015:5, 2015.

Susan G Orel, Nicole Kay, Carol Reynolds, and Daniel C Sullivan. Bi-rads categorization as a predictor of malignancy. *Radiology*, 211(3):845–850, 1999.

Serhat Özekes, Onur Osman, and A Yilmaz Çamurcu. Mammographic mass detection using a mass template. *Korean Journal of Radiology*, 6(4):221–228, 2005.

Athanasios Papadopoulos, Dimitrios I Fotiadis, and Aristidis Likas. Characterization of clustered microcalcifications in digitized mammograms using neural networks and support vector machines. *Artificial Intelligence in Medicine*, 34(2):141–150, 2005.

Athanasios Papadopoulos, Dimitrios I Fotiadis, and Lena Costaridou. Improvement of microcalcification cluster detection in mammography utilizing image enhancement techniques. *Computers in Biology and Medicine*, 38(10):1045–1055, 2008.

A Papaevangelou, S Chatzistergos, KS Nikita, and G Zografos. Breast density: computerized analysis on digitized mammograms. *Hellenic Journal of Surgery*, 83 (3):133–138, 2011.

Evgenia Papavasileiou, Redona Brahimetaj, Frederik Temmermans, Bruno Cornelis, Inneke Willekens, Johan De Mey, and Bart Jansen. Towards a cad system for breast cancer based on individual microcalcifications? In *2018 IEEE 20th International*

*Conference on e-Health Networking, Applications and Services (Healthcom)*, pages 1–5. IEEE, 2018.

JM Park, HK Choi, S-J Bae, M-S Lee, S-H Ahn, and G Gong. Clustering of breast microcalcifications: revisited. *Clinical radiology*, 55(2):114–118, 2000.

D Max Parkin, Freddie Bray, J Ferlay, and Paola Pisani. Global cancer statistics, 2002. *CA: A Cancer Journal for Clinicians*, 55(2):74–108, 2005.

D Maxwell Parkin and Leticia MG Fernández. Use of statistics to assess the global burden of breast cancer. *The Breast Journal*, 12:S70–S80, 2006.

Ana C Patrocinio, Homero Schiabel, and Romero Roseli Ap F. Evaluation of bayesian network to classify clustered microcalcifications. In *Medical Imaging 2004: Image Processing*, volume 5370, pages 1026–1034. International Society for Optics and Photonics, 2004.

Alex P Pentland. Fractal-based description of natural scenes. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (6):661–674, 1984.

Amnart Petpon and Sanun Srisuk. Face recognition with local line binary pattern. In *2009 Fifth International Conference on Image and Graphics*, pages 533–539. IEEE, 2009.

Styliani Petroudi, Timor Kadir, and Michael Brady. Automatic classification of mammographic parenchymal patterns: A statistical approach. In *Engineering in Medicine and Biology Society*, volume 1, pages 798–801, 2003.

Manju Bala Popli, Rahul Teotia, Meenakshi Narang, and Hare Krishna. Breast positioning during mammography: mistakes to be avoided. *Breast cancer: Basic and Cinical Research*, 8:BCBCR–S17617, 2014.

N Prasad, Dana Houserkova, et al. The role of various modalities in breast imaging. *Biomedical Papers of the Medical Faculty of Palacky University in Olomouc*, 151 (2), 2007.

Mellisa Pratiwi, Jeklin Harefa, Sakka Nanda, et al. Mammograms classification using gray-level co-occurrence matrix and radial basis function neural network. *Procedia Computer Science*, 59:83–91, 2015.

Amir Qaseem, Vincenza Snow, Katherine Sherif, Mark Aronson, Kevin B Weiss, and Douglas K Owens. Screening mammography for women 40 to 49 years of age: a clinical practice guideline from the american college of physicians. *Annals of Internal Medicine*, 146(7):511–515, 2007.

Yanpeng Qu, Changjing Shang, Wei Wu, and Qiang Shen. Evolutionary fuzzy extreme learning machine for mammographic risk analysis. *International Journal of Fuzzy Systems*, 13(4), 2011.

Farshid Rafiee Rad, Hamid Soltanian-Zadeh, Mohammad Rahmati, and Syamak Pour-Abdollah. Microcalcification classification in mammograms using multiwavelet features. In *Wavelet Applications in Signal and Image Processing VII*, volume 3813, pages 832–842. International Society for Optics and Photonics, 1999.

Andrik Rampun, Philip Morrow, Bryan Scotney, and John Winder. Breast density classification using local ternary patterns in mammograms. In *International Conference Image Analysis and Recognition*, pages 463–470. Springer, 2017.

Andrik Rampun, Bryan William Scotney, Philip John Morrow, Hui Wang, and John Winder. Breast density classification using local quinary patterns with various neighbourhood topologies. *Journal of Imaging*, 4(1):14, 2018a.

Andrik Rampun, Hui Wang, Bryan Scotney, Philip Morrow, and Reyer Zwiggelaar. Classification of mammographic microcalcification clusters with machine learning confidence levels. In *14th International Workshop on Breast Imaging (IWBI 2018)*, volume 10718, page 107181B. International Society for Optics and Photonics, 2018b.

A Redondo, M Comas, F Macia, F Ferrer, C Murta-Nascimento, MT Maristany, E Molins, M Sala, and X Castells. Inter-and intraradiologist variability in the bi-rads assessment and breast density categories for screening mammograms. *The British Journal of Radiology*, 85(1019):1465–1470, 2012.

Jinchang Ren. Ann vs. svm: Which one performs better in classification of mccs in mammogram imaging. *Knowledge-Based Systems*, 26:144–153, 2012.

Breast Imaging Reporting. Data system, bi-rads: Mammography. *Reston: Am College Radiology*, 2003.

Olivier Rioul and Martin Vetterli. Wavelets and signal processing. *IEEE Signal Processing Magazine*, 8(4):14–38, 1991.

Lawrence G Roberts. *Machine perception of three-dimensional solids*. PhD thesis, Massachusetts Institute of Technology, 1963.

Galina L Rogova, Paul C Stomper, and Chih-Chung Ke. Microcalcification texture analysis in a hybrid system for computer-aided mammography. In *Medical Imaging 1999: Image Processing*, volume 3661, pages 1426–1434. International Society for Optics and Photonics, 1999.

JP Rolland and Harrison H Barrett. Effect of random background inhomogeneity on observer detection performance. *JOSA A*, 9(5):649–658, 1992.

Cristina Romero, Asunción Almenar, Jose María Pinto, Celia Varela, Enriqueta Muñoz, and Miguel Botella. Impact on breast cancer diagnosis in a multidisciplinary unit after the incorporation of mammography digitalization and computer-aided detection systems. *American Journal of Roentgenology*, 197(6):1492–1497, 2011.

Chris Rose, Daniele Turi, Alan Williams, Katy Wolstencroft, and Chris Taylor. Web services for the ddsm and digital mammography research. In *International workshop on digital mammography*, pages 376–383. Springer, 2006.

CJ Rose and CJ Taylor. A statistical model of texture for medical image synthesis and analysis. *Med. Image Understand. Anal*, pages 1–4, 2003.

Robert D Rosenberg, William C Hunt, Michael R Williamson, Frank D Gilliland, Philip W Wiest, Charles A Kelsey, Charles R Key, and Michael N Linver. Effects of age, breast density, ethnicity, and estrogen replacement therapy on screening mammographic sensitivity and cancer stage at diagnosis: review of 183,134 screening mammograms in albuquerque, new mexico. *Radiology*, 209(2):511–518, 1998.

Azriel Rosenfeld. Visual texture analysis: an overview. Technical report, Maryland Univ College Park Computer Science Center, 1975.

Seng Roty, Cholatip Wiratkapun, Rawesak Tanawongsuwan, and Sukanya Phongsuphap. Analysis of microcalcification features for pathological classification of mammograms. In *2017 10th Biomedical Engineering International Conference (BMEiCON)*, pages 1–5. IEEE, 2017.

Daniel B Russakoff and Akira Hasegawa. Generation and application of a probabilistic breast cancer atlas. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 454–461. Springer, 2006.

Audrey F Saftlas and Moyses Szklo. Mammographic parenchymal patterns and breast cancer risk. *Epidemiologic Reviews*, 9(1):146–174, 1987.

Punam K Saha, Jayaram K Udupa, Emily F Conant, Dev P Chakraborty, and Daniel Sullivan. Breast tissue density quantification via digitized mammograms. *IEEE Transactions on Medical Imaging*, 20(8):792–803, 2001.

Nafiza Saidin, Harsa Amylia Mat Sakim, Umi Kalthum Ngah, and Ibrahim Lutfi Shuaib. Segmentation of breast regions in mammogram based on density: a review. *arXiv preprint arXiv:1209.5494*, 2012.

Soudeh Saien, Abdol Hamid Pilevar, and Hamid Abrishami Moghaddam. Refinement of lung nodule candidates based on local geometric shape analysis and laplacian of gaussian kernels. *Computers in Biology and Medicine*, 54:188–198, 2014.

Shelda Sajeev, Mariusz Bajger, and Gobert Lee. Superpixel pattern graphs for identifying breast mass rois in dense background: a preliminary study. In *14th International Workshop on Breast Imaging (IWBI 2018)*, volume 10718, page 107180V. International Society for Optics and Photonics, 2018.

E Sakka, A Prentza, and D Koutsouris. Classification algorithms for microcalcifications in mammograms. *Oncology Reports*, 15(4):1049–1055, 2006.

E Sala, L Solomon, R Warren, J McCann, S Duffy, R Luben, and N Day. Size, node status and grade of breast tumours: association with mammographic parenchymal patterns. *European Radiology*, 10(1):157–161, 2000.

Ravi K Samala, Heang-Ping Chan, Lubomir M Hadjiiski, Kenny Cha, and Mark A Helvie. Deep-learning convolution neural network for computer-aided detection of microcalcifications in digital breast tomosynthesis. In *Medical Imaging 2016: Computer-Aided Diagnosis*, volume 9785, page 97850Y. International Society for Optics and Photonics, 2016.

Alexander Samborskiy. Cell detection and counting for microscope images with applications to stem cell engineering. *Department of Electrical and Computer Engineering, Pittsburgh*, 2007.

Nirupam Sarkar and BB Chaudhuri. An efficient differential box-counting approach to compute fractal dimension of image. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(1):115–120, 1994.

Ferdinand Schmidt, Erich Sorantin, Csaba Szepesvari, Ewald Graif, Michael Becker, Heinz Mayer, and Karin Hartwagner. An automatic method for the identification and interpretation of clustered microcalcifications in mammograms. *Physics in Medicine and Biology*, 44(5):1231–1244, 1999.

S Sentelle, C Sentelle, and Melanie A Sutton. Multiresolution-based segmentation of calcifications for the early detection of breast cancer. *Real-Time Imaging*, 8(3): 237–252, 2002.

Mehmet Sezgin and Bülent Sankur. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1):146–166, 2004.

Yaniv Shachor, Hayit Greenspan, and Jacob Goldberger. A mixture of views network with applications to the classification of breast microcalcifications. *arXiv preprint arXiv:1803.06898*, 2018.

Yuan-Zhi Shao, Li-Zhi Liu, Meng-Jie Bie, Chan-chan Li, Yao-pan Wu, Xiao-ming Xie, and Li Li. Characterizing the clustered microcalcifications on mammograms to predict the pathological classification and grading: A mathematical modeling approach. *Journal of Digital Imaging*, 24(5):764, 2011.

Neeraj Sharma, Amit K Ray, Shiru Sharma, KK Shukla, Satyajit Pradhan, and Lalit M Aggarwal. Segmentation and classification of medical images using texture-primitive features: Application of bam-type artificial neural network. *Journal of Medical Physics/Association of Medical Physicists of India*, 33(3):119, 2008.

Vipul Sharma. Comparative analysis of region of interest of different sizes for breast density classification. *International Journal of Medical Research  Health Sciences*, 6:76–84, 2017.

Liang Shen, Rangaraj M Rangayyan, and JE Leo Desautels. Detection and classification of mammographic calcifications. *International Journal of Pattern Recognition and Artificial Intelligence*, 7(06):1403–1416, 1993.

Liang Shen, Rangaraj M Rangayyan, and JE Leo Desautels. Application of shape analysis to mammographic calcifications. *IEEE Transactions on Medical Imaging*, 13(2):263–274, 1994.

Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *Departmental Papers (CIS)*, page 107, 2000.

EA Sickles, CJ d'Orsi, LW Bassett, CM Appleton, WA Berg, ES Burnside, et al. Acr bi-rads® mammography. *ACR BI-RADS® Atlas, Breast imaging reporting and data system*, 5, 2013.

Edward A Sickles. Breast calcifications: mammographic evaluation. *Radiology*, 160 (2):289–293, 1986.

Edward A Sickles. Practical solutions to common mammographic problems: tailoring the examination. *American Journal of Roentgenology*, 151(1):31–39, 1988.

Edward A Sickles. Wolfe mammographic parenchymal patterns and breast cancer risk. *American Journal of Roentgenology*, 188(2):301–303, 2007.

Birmohan Singh and Manpreet Kaur. An approach for enhancement of microcalcifications in mammograms. *Journal of Medical and Biological Engineering*, 37(4): 567–579, 2017.

Sameer Singh and Keir Bovis. A weighted gaussian mixture model with markov random fields and adaptive expert combination strategy for segmenting masses in mammograms. *Chapter*, 8:263–289, 2006.

Sukhwinder Singh, Vinod Kumar, HK Verma, and Dilbag Singh. Svm based system for classification of microcalcifications in digital mammograms. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4747–4750. IEEE, 2006.

S Eva Singletary. Rating the risk factors for breast cancer. *Annals of surgery*, 237(4): 474, 2003.

Radhika Sivaramakrishna and Richard Gordon. Detection of breast cancer at a smaller size can reduce the likelihood of metastatic spread: a quantitative analysis. *Academic Radiology*, 4(1):8–12, 1997.

Radhika Sivaramakrishna, Nancy A Obuchowski, William A Chilcote, and Kimerly A Powell. Automatic segmentation of mammographic density. *Academic Radiology*, 8(3):250–256, 2001.

Jack Sklansky. Image segmentation and feature extraction. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(4):237–247, 1978.

Pierre Soille. *Morphological image analysis: principles and applications*. Springer Science & Business Media, 2013.

Hamid Soltanian-Zadeh, Siamak Pourabdollah-Nezhad, and Farshid Rafiee Rad. Shape-based and texture-based feature extraction for classification of microcalcifications in mammograms. In *Medical Imaging 2001: Image Processing*, volume 4322, pages 301–311. International Society for Optics and Photonics, 2001.

Hamid Soltanian-Zadeh, Farshid Rafiee-Rad, et al. Comparison of multiwavelet, wavelet, haralick, and shape features for microcalcification classification in mammograms. *Pattern Recognition*, 37(10):1973–1986, 2004.

Mary Scott Soo, Eric L Rosen, Jessie Q Xia, Sujata Ghate, and Jay A Baker. Computer-aided detection of amorphous calcifications. *American Journal of Roentgenology*, 184(3):887–892, 2005.

GN Srinivasan and G Shobha. Statistical texture analysis. In *Proceedings of world academy of Science, Engineering and Technology*, volume 36, pages 1264–1269, 2008.

Katarzyna Stąpor. Evaluating and comparing classifiers: Review, some recommendations and limitations. In *International Conference on Computer Recognition Systems*, pages 12–21. Springer, 2017.

Elizabeth Steiner, David Klubert, and Doug Knutson. Assessing breast cancer risk in women. *Am Fam Physician*, 78(12):1361–6, 2008.

Tomislav Stojić and Branimir Reljin. Enhancement of microcalcifications in digitized mammograms: Multifractal and mathematical morphology approach. *FME Transactions*, 38(1):1–9, 2010.

Harry Strange, Zhili Chen, Erika RE Denton, and Reyer Zwiggelaar. Modelling mammographic microcalcification clusters using persistent mereotopology. *Pattern Recognition Letters*, 47:157–163, 2014.

TS Subashini, Vennila Ramalingam, and S Palanivel. Automated assessment of breast tissue density in digital mammograms. *Computer Vision and Image Understanding*, 114(1):33–43, 2010.

John Suckling, J Parker, D Dance, S Astley, I Hutt, C Boggis, I Ricketts, E Stamatakis, N Cerneaz, S Kok, et al. The mammographic image analysis society digital mammogram database. In *Exerpta Medica. International Congress Series*, volume 1069, pages 375–378, 1994.

Zobia Suhail, Erika RE Denton, and Reyer Zwiggelaar. Classification of microcalcification in mammograms using scalable linear fisher discriminant analysis. *Medical & Biological Engineering & Computing*, 56(8):1475–1485, 2018a.

Zobia Suhail, Erika RE Denton, and Reyer Zwiggelaar. Tree-based modelling for the classification of mammographic benign and malignant micro-calcification clusters. *Multimedia Tools and Applications*, 77(5):6135–6148, 2018b.

Dixit G Swati. Improved algorithm for blob detection in document images. In *Proceedings of the 2014 5th International Conference Confluence The Next Generation Information Technology Summit (Confluence 2014)*, pages 703–708, 2014.

L Tabár, T Tot, PB Dean, and B Cancer. The art and science of early detection with mamography: Perception, interpretation, histopatholigic correlation, 2004.

László Tabár and Peter B Dean. Mammographic parenchymal patterns: risk indicator for breast cancer? *Jama*, 247(2):185–189, 1982.

Laszlo Tabar, Ming-Fang Yen, Bedrich Vitak, Hsiu-Hsi Tony Chen, Robert A Smith, and Stephen W Duffy. Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. *The Lancet*, 361(9367):1405–1410, 2003.

Hideyuki Tamura, Shunji Mori, and Takashi Yamawaki. Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, 8 (6):460–473, 1978.

Maxine Tan, Bin Zheng, Pandiyarajan Ramalingam, and David Gur. Prediction of near-term breast cancer risk based on bilateral mammographic feature asymmetry. *Academic Radiology*, 20(12):1542–1550, 2013.

Nutao Tan, Lei Huang, and Changping Liu. A new probabilistic local binary pattern for face verification. In *2009 16th IEEE International Conference on Image Processing (ICIP)*, 2009.

Xiaoyang Tan and Bill Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing*, 19(6): 1635–1650, 2010.

CG Taylor, J Champness, M Reddy, P Taylor, HWW Potts, and R Given-Wilson. Reproducibility of prompts in computer-aided detection (cad) of breast cancer. *Clinical Radiology*, 58(9):733–738, 2003.

K Taylor, P Britton, S O'keeffe, and MG Wallis. Quantification of the uk 5-point breast imaging classification and mapping to bi-rads to facilitate comparison with international literature. *The British Journal of Radiology*, 84(1007):1005–1010, 2011.

Paul Taylor, Robyn Owens, and David Ingram. Simulated mammography using synthetic 3d breasts. In *Digital Mammography*, pages 283–290. Springer, 1998.

Richard P Taylor, Adam P Micolich, and David Jonas. Fractal expressionism. *Physics World*, 12(10):25, 1999.

Philip M Tchou, Tamara Miner Haygood, E Neely Atkinson, Tanya W Stephens, Paul L Davis, Elsa M Arribas, William R Geiser, and Gary J Whitman. Interpretation time of computer-aided detection at screening mammography. *Radiology*, 257(1):40–46, 2010.

David L Thiele, Carolyn Kimme-Smith, Timothy D Johnson, Marie McCombs, and Lawrence W Bassett. Using tissue texture surrounding calcification clusters to predict benign vs malignant outcomes. *Medical Physics*, 23(4):549–555, 1996.

Andy Tirtajaya and Diaz D Santika. Classification of microcalcification using dual-tree complex wavelet transform and support vector machine. In *2010 Second International Conference on Advances in Computing, Control, and Telecommunication Technologies*, pages 164–166. IEEE, 2010.

Linda Titus-Ernstoff, Anna NA Tosteson, Claudia Kasales, Julia Weiss, Martha Goodrich, Elizabeth E Hatch, and Patricia A Carney. Breast cancer risk factors in relation to breast density (united states). *Cancer Causes & Control*, 17(10): 1281–1290, 2006.

Hector Tobon and Hernando Salazar. Ultrastructure of the human mammary gland. ii. postpartum lactogenesis. *The Journal of Clinical Endocrinology & Metabolism*, 40 (5):834–844, 1975.

Lindsey A Torre, Freddie Bray, Rebecca L Siegel, Jacques Ferlay, Joannie Lortet-Tieulent, and Ahmedin Jemal. Global cancer statistics, 2012. *CA: A Cancer Journal for Clinicians*, 65(2):87–108, 2015.

Albert Torrent, Anton Bardera, Arnau Oliver, Jordi Freixenet, Imma Boada, Miguel Feixes, Robert Martí, Xavier Lladó, Josep Pont, Elsa Pérez, et al. Breast density segmentation: a comparison of clustering and region based techniques. In *International Workshop on Digital Mammography*, pages 9–16. Springer, 2008.

Gabriela Torres-Mejía, Bianca De Stavola, Diane S Allen, Juan J Pérez-Gavilán, Jorge M Ferreira, Ian S Fentiman, and Isabel dos Santos Silva. Mammographic features and subsequent risk of breast cancer: a comparison of qualitative and quantitative evaluations in the guernsey prospective studies. *Cancer Epidemiology and Prevention Biomarkers*, 14(5):1052–1059, 2005.

Meritxell Tortajada, Arnau Oliver, Robert Martí, Mariona Vilagran, Sergi Ganau, Lidia Tortajada, Melcior Sentís, and Jordi Freixenet. Adapting breast density classification from digitized to full-field digital mammograms. In *International Workshop on Digital Mammography*, pages 561–568. Springer, 2012.

Gerard V Trunk. A problem of dimensionality: A simple example. *IEEE Transactions on pattern analysis and machine intelligence*, (3):306–307, 1979.

Mihran Tuceryan. Moment-based texture segmentation. *Pattern Recognition Letters*, 15(7):659–668, 1994.

Mihran Tuceryan and Anil K. Jain. Texture segmentation using voronoi polygons. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(2):211–216, 1990.

Mihran Tuceryan and Anil K Jain. Texture analysis. In *Handbook of pattern recognition and computer vision*, pages 235–276. World Scientific, 1993.

Stylianos D Tzikopoulos, Michael E Mavroforakis, Harris V Georgiou, Nikos Dimitropoulos, and Sergios Theodoridis. A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry. *Computer Methods and Programs in Biomedicine*, 102(1):47–63, 2011.

Devrim Unay, Ahmet Ekin, Mujdat Cetin, Radu Jasinschi, and Aytul Ercil. Robustness of local binary patterns in brain mr image analysis. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 2098–2101. IEEE, 2007.

Michael Unser. Texture classification and segmentation using wavelet frames. *IEEE Transactions on Image Processing*, 4(11):1549–1560, 1995.

Noelia Vállez, Gloria Bueno, Oscar Déniz-Suárez, José A Seone, Julián Dorado, and Alejandro Pazos. A tree classifier for automatic breast tissue classification based on birads categories. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 580–587. Springer, 2011.

Manik Varma and Andrew Zisserman. A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62(1-2):61–81, 2005.

Manik Varma and Andrew Zisserman. A statistical approach to material classification using image patch exemplars. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(11):2032–2047, 2009.

Wouter JH Veldkamp, Nico Karssemeijer, Johannes DM Otten, and Jan HCL Hendriks. Automated classification of clustered microcalcifications into malignant and benign types. *Medical Physics*, 27(11):2600–2608, 2000.

Wouter JH Veldkamp, Nico Karssemeijer, and Jan HCL Hendriks. Experiments with radiologists and a fully automated method for characterization of microcalcification clusters. In *International Congress Series*, volume 1230, pages 586–592. Elsevier, 2001.

J Virmani and V Kumar. Quantitative evaluation of image enhancement techniques. In *Proceedings of International Conference on Biomedical Engineering and Assistive Technology (BEATS)*, pages 1–8, 2010.

Jitendra Virmani, Nilanjan Dey, Vinod Kumar, et al. Pca-pnn and pca-svm based cad systems for breast density classification. In *Applications of intelligent optimization in biology and medicine*, pages 159–180. Springer, 2016.

Sunhua Wan, Hsiang-Chieh Lee, Xiaolei Huang, Ting Xu, Tao Xu, Xianxu Zeng, Zhan Zhang, Yuri Sheikine, James L Connolly, James G Fujimoto, et al. Integrated local binary pattern texture features for classification of breast tissue imaged by optical coherence microscopy. *Medical Image Analysis*, 38:104–116, 2017.

Jinhua Wang, Xi Yang, Hongmin Cai, Wanchang Tan, Cangzheng Jin, and Li Li. Discrimination of breast cancer with microcalcifications on mammography by deep learning. *Scientific Reports*, 6:27327, 2016.

Xingwei Wang, Dror Lederman, Jun Tan, Xiao Hui Wang, and Bin Zheng. Computerized detection of breast tissue asymmetry depicted on bilateral mammograms: a preliminary study of breast risk stratification. *Academic Radiology*, 17(10):1234–1241, 2010.

Zhi-Zhong Wang and Jun-Hai Yong. Texture analysis and classification with linear regression model based on wavelet transform. *IEEE Transactions on Image Processing*, 17(8):1421–1430, 2008.

Simon K Warfield, Kelly H Zou, and William M Wells. Validation of image segmentation by estimating rater bias and variance. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 839–847. Springer, 2006.

Linda J Warren Burhenne, Susan A Wood, Carl J D'Orsi, Stephen A Feig, Daniel B Kopans, Kathryn F O'Shaughnessy, Edward A Sickles, Laszlo Tabar, Carl J Vyborny, and Ronald A Castellino. Potential contribution of computer-aided detection to the sensitivity of screening mammography. *Radiology*, 215(2):554–562, 2000.

Jun Wei, Songfeng Li, Heang-Ping Chan, Mark A Helvie, Marilyn A Roubidoux, Yao Lu, Chuan Zhou, Lubomir Hadjiiski, and Ravi K Samala. Deep convolutional neural network for mammographic density segmentation. In *Medical Imaging 2018: Computer-Aided Diagnosis*, volume 10575, page 105753I. International Society for Optics and Photonics, 2018.

Liyang Wei, Yongyi Yang, Robert M Nishikawa, and Yulei Jiang. A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications. *IEEE Transactions on Medical Imaging*, 24(3):371–380, 2005a.

Liyang Wei, Yongyi Yang, Robert M Nishikawa, Miles N Wernick, and Alexandra Edwards. Relevance vector machine for automatic detection of clustered microcalcifications. *IEEE Transactions on Medical Imaging*, 24(10):1278–1285, 2005b.

Liyang Wei, Yongyi Yang, and Robert M Nishikawa. Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis. *Pattern recognition*, 42(6):1126–1132, 2009.

H Gilbert Welch and Brittney A Frankel. Likelihood that a woman with screen-detected breast cancer has had her "life saved" by that screening. *Archives of Internal Medicine*, 171(22):2043–2046, 2011.

Joan S Weszka, Charles R Dyer, and Azriel Rosenfeld. A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man, and Cybernetics*, (4):269–285, 1976.

Peter Winkler, Ewald Graif, Csaba Szepesvári, Michael Becker, Heinz Mayer, Ferdinand Schmidt, and Erich Sorantin. Computer aided diagnosis of clustered microcalcifications using artificial neural nets. *Journal of Computing and Information Technology*, 8(2):151–160, 2000.

Fred Winsberg, Milton Elkin, Josiah Macy Jr, Victoria Bordaz, and William Weymouth. Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis. *Radiology*, 89(2):211–215, 1967.

Andrew P Witkin. Scale-space filtering. In *Readings in Computer Vision*, pages 329–332. Elsevier, 1987.

Ian H Witten and Eibe Frank. Data mining: practical machine learning tools and techniques with java implementations. *Acm Sigmod Record*, 31(1):76–77, 2002.

Brendt Wohlberg and Gerhard De Jager. A review of the fractal image coding literature. *IEEE Transactions on Image Processing*, 8(12):1716–1729, 1999.

John N Wolfe. Breast patterns as an index of risk for developing breast cancer. *American Journal of Roentgenology*, 126(6):1130–1137, 1976a.

John N Wolfe. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer*, 37(5):2486–2492, 1976b.

Kevin S Woods, Christopher C Doss, Kevin W Bowyer, Jeffrey L Solka, Carey E Priebe, and W PHILIP KEGELMEYER JR. Comparative evaluation of pattern recognition techniques for detection of microcalcifications in mammography. *International Journal of Pattern Recognition and Artificial Intelligence*, 7(06):1417–1436, 1993.

Lide Wu and Zhaohui Xie. Scaling theorems for zero-crossings. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (1):46–54, 1990.

Yong Xia, Dagan Feng, and Rongchun Zhao. Morphology-based multifractal estimation for texture segmentation. *IEEE Transactions on Image Processing*, 15(3): 614–623, 2006.

Huang Xiaohua, Su-Jing Wang, Xin Liu, Guoying Zhao, Xiaoyi Feng, and Matti Pietikainen. Discriminative spatiotemporal local binary pattern with revisited integral projection for spontaneous facial micro-expression recognition. *IEEE Transactions on Affective Computing*, 2017.

Xianchuan Xu and Qi Zhang. Medical image retrieval using local binary patterns with image euclidean distance. *2009 International Conference on Information Engineering and Computer Science*, pages 1–4, 2009.

Yong Xu, Hui Ji, and Cornelia Fermuller. A projective invariant for textures. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 1932–1939. IEEE, 2006.

E Yilmaz, B Lebe, P Balci, S Sal, and T Canda. Comparison of mammographic and sonographic findings in typical and atypical medullary carcinomas of the breast. *Clinical Radiology*, 57(7):640–645, 2002.

Siamak Yousefi and Nasser Kehtarnavaz. A new stochastic image model based on markov random fields and its application to texture modeling. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 1285–1288. IEEE, 2011.

Songyang Yu and Ling Guan. A cad system for the automatic detection of clustered microcalcifications in digitized mammogram films. *IEEE Transactions on Medical Imaging*, 19(2):115–126, 2000.

Jing-Hua Yuan, Hao-Dong Zhu, Yong Gan, and Li Shang. Enhanced local ternary pattern for texture classification. In *International Conference on Intelligent Computing*, pages 443–448. Springer, 2014.

Cauê Zaghetto, Mateus Mendelson, Alexandre Zaghetto, and Flávio de B Vidal. Liveness detection on touchless fingerprint devices using texture descriptors and artificial neural networks. In *Biometrics (IJCB), 2017 IEEE International Joint Conference on*, pages 406–412. IEEE, 2017.

Jaromir Zavadil, Jiri Tuma, and Vitor MF Santos. Traffic signs detection using blob analysis and pattern recognition. In *Proceedings of the 13th International Carpathian Control Conference (ICCC)*, pages 776–779. IEEE, 2012.

Annika Zdon, Mark A Helvie, Alex Tsodikov, Heang-Ping Chan, and Jun Wei. Analysis of mammographic density as a predictor for breast cancer occurrence. In *Medical Imaging 2019: Imaging Informatics for Healthcare, Research, and Applications*, volume 10954, page 1095412. International Society for Optics and Photonics, 2019.

Yiqiang Zhan, Yangming Ou, Michael Feldman, John Tomaszeweski, Christos Davatzikos, and Dinggang Shen. Registering histologic and mr images of prostate for image-based cancer detection. *Academic Radiology*, 14(11):1367–1381, 2007.

Baochang Zhang, Yongsheng Gao, Sanqiang Zhao, and Jianzhuang Liu. Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor. *IEEE Transactions on Image Processing*, 19(2):533–544, 2010.

Jianguo Zhang and Tieniu Tan. Brief review of invariant texture analysis methods. *Pattern Recognition*, 35(3):735–747, 2002.

Wenchao Zhang, Shiguang Shan, Wen Gao, Xilin Chen, and Hongming Zhang. Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 786–791. IEEE, 2005.

Chuan Zhou, Heang-Ping Chan, Nicholas Petrick, Mark A Helvie, Mitchell M Goodsitt, Berkman Sahiner, and Lubomir M Hadjiiski. Computerized image analysis: estimation of breast density on mammograms. *Medical Physics*, 28(6):1056–1069, 2001.

Dongxiao Zhou. *Texture analysis and synthesis using a generic Markov-Gibbs image model*. PhD thesis, ResearchSpace@ Auckland, 2006.

Gustavo Zucca-Matthes, Cícero Urban, and André Vallejo. Anatomy of the nipple and breast ducts. *Gland Surgery*, 5(1):32, 2016.

R Zwiggelaar, I Muhimmah, and ERE Denton. Mammographic density classification based on statistical gray-level histogram modelling. In *Proc. Medical Image Understanding and Analysis*, pages 183–186, 2005.

Reyer Zwiggelaar. Local greylevel appearance histogram based texture segmentation. In *International Workshop on Digital Mammography*, pages 175–182. Springer, 2010.

Imad Zyout and Ikhlas Abdel-Qader. Classification of microcalcification clusters via pso-knn heuristic parameter selection and glcm features. *International Journal of Computer Applications*, 31(2):34–39, 2011.

Imad Zyout, Ikhlas Abdel-Qader, and Christina Jacobs. Bayesian classifier with simplified learning phase for detecting microcalcifications in digital mammograms. *Journal of Biomedical Imaging*, 2009:30, 2009.